



Catarina Maria Couto Amorim Marques

Mestrado em Engenharia Biomédica

**Descodificação visual de expressões faciais
através de um *eye tracker* de baixo custo**

Dissertação para obtenção do Grau de Mestre em
Engenharia Biomédica

Orientador: Ricardo Nuno Pereira Verga e Afonso Vigário, Professor Associado, Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa

Co-orientadora: Ana Patrícia Teixeira Pinheiro, Professora Auxiliar, Faculdade de Psicologia da Universidade de Lisboa



FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Junho, 2020

Descodificação visual de expressões faciais através de um *eye tracker* de baixo custo

Copyright © Catarina Maria Couto Amorim Marques, Faculdade de Ciências e Tecnologia, Universidade NOVA de Lisboa.

A Faculdade de Ciências e Tecnologia e a Universidade NOVA de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objetivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

AGRADECIMENTOS

Quero agradecer a todas as pessoas que ficaram ao meu lado durante estes memoráveis anos de estudo, diversão, crescimento pessoal e profissional. Em particular, quero agradecer a quem me apoiou no desenvolvimento desta dissertação.

Em primeiro lugar, quero agradecer ao professor e orientador Ricardo Vigário por me ter dado a oportunidade de desenvolver este projeto e por todo o acompanhamento e sensibilidade demonstrada durante todo este processo. Quero agradecer à professora e co-orientadora Ana Patrícia Pinheiro pelo cuidado e apoio no desenvolvimento deste trabalho..

Para todos os meus grandes amigos, um grande obrigado. Convosco, estes cinco anos de trabalho foram muito mais fáceis e cheios de companhia e diversão! Quero agradecer em particular aos meus amigos Miguel Andrade, Bernardo Teixeira e Carlos Mendes, pois sem eles o meu percurso nesta faculdade não teria sido o mesmo. Um enorme obrigado também à minha colega de casa Francisca que me acompanhou e me motivou muito durante a fase final deste trabalho.

Por fim, estou muito grata pelos meus pais, os meus avós e os meus irmãos que me apoiaram sempre e nunca deixam de acreditar em mim.

RESUMO

As expressões faciais são uma forma de comunicação não-verbal muito importante para a transmissão de informações sociais entre seres-humanos e a sua descodificação é um processo eminentemente automático e espontâneo. Existem fatores que podem fazer variar a atenção visual dada a cada elemento da expressão facial, podendo esta atenção depender da emoção apresentada ou do tipo de observador.

A monitorização da atenção visual durante este processo poderá ser feita através de um *eye tracker*. Porém, estes dispositivos são muito dispendiosos. Para estudar a dinâmica dos movimentos oculares durante a descodificação de expressões faciais com diferentes cargas emocionais, recorreu-se a um *eye tracker* de baixo custo, desenvolvido e otimizado em teses anteriores [1, 2].

Sendo este o primeiro estudo realizado com o referido dispositivo em contexto real, foi primeiramente necessário validá-lo. Para tal, seguiu-se a metodologia de um estudo realizado em 2011[3] cujo objetivo foi examinar a evolução da direção do olhar de indivíduos saudáveis enquanto estes examinam diferentes expressões faciais representando emoções básicas e investigar se existem padrões, na dinâmica de atenção visual, específicos a cada emoção apresentada. Depois de validado o dispositivo, foi realizado um estudo comparativo dos movimentos oculares, durante a descodificação de expressões faciais emocionais, entre indivíduos com formação musical e indivíduos sem este tipo de formação.

De um modo geral observou-se um foco da atenção visual nos olhos, independentemente da emoção apresentada. A região da boca revelou ser um alvo mais frequente da atenção visual na emoção alegria em comparação com as outras emoções. Para o grupo de indivíduos com formação musical, verificou-se uma frequência e duração de fixações na região dos olhos superior aos valores apresentados para o grupo sem formação musical e foi também verificada uma dinâmica de olhar menos dispersa e menos errática para o grupo com formação musical.

Palavras-chave: *eye tracking*, atenção visual, emoção, descodificação facial, dinâmica visual, formação musical

ABSTRACT

Facial expressions are a very important form of non-verbal communication for the transmission of social information between human beings and its decoding is an eminently automatic and spontaneous process. There are factors that can impact the visual attention given to each element of the facial expression, and this attention may depend on the emotion presented or the type of observer.

Monitoring of visual attention during this process can be done using an *eye tracker*. However, these are very expensive. To study the dynamics of eye movements during the decoding of emotional expressions, a low-cost *eye tracker* was developed and optimized in previous theses [1, 2].

Since this is the first study carried out with the referred device in a real context, it was first necessary to validate it. To this end, we followed the methodology of a study carried out in 2011 [3] whose objective was to examine the scanpaths of healthy individuals while they observed different facial expressions representing basic emotions. After validating the device, a comparative study of eye movements was performed, during the decoding of emotional expressions, between individuals with musical training and individuals without this type of training.

In general, there was a focus of visual attention in the eyes, regardless of the emotion presented. The mouth region proved to be a more frequent target of visual attention in the emotion of joy compared to other emotions. For the group of individuals with musical training, there was a frequency and duration of fixations in the eye region higher than the values presented for the group without musical training and there was also a less dispersed and less erratic scanpath for the group with musical training.

Keywords: visual attention, emotion, facial decoding, scanpath, musical training

ÍNDICE

Lista de Figuras	xiii
Lista de Tabelas	xv
1 Introdução	1
1.1 Contextualização	1
1.2 Objetivos	3
2 Conceitos Teóricos	5
2.1 O cérebro humano	5
2.1.1 Percepção e formação musical	7
2.1.2 Decodificação de expressões faciais emocionais	7
2.2 <i>Eye tracking</i>	9
2.2.1 Movimentos oculares	9
2.2.2 Ângulo Visual	10
2.2.3 Sistema de <i>eye tracking</i>	11
3 Estado da Arte	15
3.1 <i>Eye tracking</i>	15
3.1.1 Eye tracker de baixo custo	17
3.2 Decodificação de expressões faciais emocionais	19
3.2.1 Hemisfério direito	20
3.2.2 Monitorização dos movimentos oculares na decodificação de expressões faciais	21
4 Materiais e Métodos	23
4.1 Materiais	23
4.1.1 Telemóvel	23
4.1.2 Suporte	23
4.1.3 Computador	24
4.1.4 <i>Karolinska Directed Emotional Faces</i>	25
4.1.5 <i>Self Assessment Manikin</i>	26
4.2 Métodos	27

4.2.1	Participantes	27
4.2.2	Recolha de dados	28
4.2.3	Pré-processamento dos dados	28
4.2.4	Análise de dados	31
5	Resultados e Discussão	35
5.1	Localização da primeira fixação	35
5.1.1	Validação geral	36
5.1.2	Efeito da formação musical	36
5.2	Duração média das fixações	39
5.2.1	Validação geral	40
5.2.2	Efeito da formação musical	40
5.3	Número médio de fixações	42
5.3.1	Validação geral	43
5.3.2	Efeito da formação musical	43
5.4	Rácio olhos-boca	44
5.4.1	Validação geral	45
5.4.2	Efeito da formação musical	45
5.5	Distância total	46
5.5.1	Validação geral	47
5.5.2	Efeito da formação musical	47
5.6	Discussão	47
6	Conclusão	51
6.1	Trabalho futuro	52
	Bibliografia	53

LISTA DE FIGURAS

2.1	Hemisférios cerebrais e corpo caloso (adapt. [13]).	5
2.2	Vista ventral do lobo temporal: amígdala e área fusiforme da face (adapt. [16]).	6
2.3	Formato de um objeto inanimado que se assemelha a uma face.	8
2.4	Estruturas do olho humano (adapt.[30]).	10
2.5	Representação do ângulo visual (adapt. [33])	11
2.6	Relatório visual: traçado de olhares (<i>gaze plot</i>) (adapt. [36]).	13
2.7	Relatório visual: <i>heatmap</i> (adapt. [36]).	13
2.8	Relatório visual: mapa de zonas sombreadas (adapt. [37]).	14
2.9	Relatório visual: <i>cluster</i> (adapt. [38]).	14
3.1	À esquerda temos a lente de contacto com o rolo magnético encapsulado e à direita o aparelho para medição dos movimentos através do campo magnético (adapt. [41]).	16
3.2	Exemplos de sinais medidos através de EOG (adapt. [42]).	16
3.3	Colocação do telemóvel dentro do suporte.	18
3.4	Ordem de apresentação dos pontos de calibração [1].	18
3.5	Utilização do dispositivo durante a apresentação do estímulo.	19
4.1	À esquerda entrada na aplicação e à direita escolha de um estímulo.	24
4.2	Diferentes perspetivas do suporte em cartão. À esquerda podemos ver o interruptor do sistema e o suporte das pilhas, ao centro podemos ver a fita e o suporte do nariz e à direita podemos ver o interior do <i>cardboard</i> iluminado (adap. [2]).	25
4.3	Interface através da qual é feito o carregamento do vídeo de calibração, do vídeo de movimentos correspondente ao estímulo e do próprio estímulo [2].	25
4.4	Expressões faciais emocionais. Da esquerda para a direita: Neutro, Tristeza, Medo, Raiva e Felicidade [62].	26
4.5	Escala utilizada para a avaliação de valência.	26
4.6	Escala utilizada para a avaliação de ativação.	26
4.7	Exemplo da apresentação dos dados recolhidos em eyetracking	28
4.8	Apresentação dos dados antes (à esquerda) e após (à direita) a translação. De cima para baixo: imagem-estímulo nº 15, nº 40 e nº 80.	30

4.9	Translação manual. À esquerda antes do processamento e à direita após o processamento	31
4.10	Translação horizontal. À esquerda antes do processamento e à direita após o processamento	31
4.11	Limites das áreas faciais de interesse.	32
5.1	Valores médios da frequência da primeira fixação para cada área de interesse relativos às diferentes emoções.	36
5.2	Valores médios da frequência da primeira fixação para cada área de interesse relativos às diferentes emoções. Em cima, participantes com formação musical e em baixo, participantes sem formação musical.	37
5.3	Duração média relativa das fixações.	39
5.4	Duração média relativa das fixações. Em cima, participantes com formação musical e em baixo, participantes sem formação musical.	41
5.5	Número médio relativo de fixações.	42
5.6	Número médio relativo de fixações. À esquerda, participantes com formação musical e à direita, participantes sem formação musical.	43
5.7	Rácio médio olhos-boca ((duração de fixação nos olhos-duração de fixação na boca)/(duração de fixação nos olhos+duração de fixação na boca)).	45
5.8	Rácio médio olhos-boca ((duração de fixação nos olhos-duração de fixação na boca)/(duração de fixação nos olhos+duração de fixação na boca), relativo a participantes com e sem formação musical	46
5.9	Distância média percorrida pelo olhar.	47
5.10	Distância média percorrida pelo olhar, relativa a participantes com e sem formação musical.	48

LISTA DE TABELAS

2.1	Classificação dos movimetos oculares (adapt [34])	12
5.1	Análise de variância - Localização da primeira fixação.	38
5.2	Teste-t - Localização da primeira fixação.	38
5.3	Análise de variância - Duração média das fixações.	40
5.4	Teste-t - Duração média das fixações.	41
5.5	Análise de variância - Número médio de fixações.	44
5.6	Teste-t - Número médio de fixações.	44
5.7	Análise de variância - Rácio olhos-boca.	46
5.8	Teste-t - Rácio olhos - boca.	46

INTRODUÇÃO

1.1 Contextualização

As expressões faciais são uma forma de comunicação não-verbal muito importante para a transmissão de informações sociais entre seres-humanos. A interpretação destas é de igual modo relevante e a sua descodificação é um processo eminentemente automático e espontâneo [4].

O hemisfério direito do cérebro tem um papel central na descodificação de expressões faciais com carga emocional[5]. Enquanto que o hemisfério esquerdo é especializado em tarefas lógicas e analíticas, o hemisfério direito assume uma importância maior em tarefas que envolvam criatividade e comunicação não-verbal [6]. Por sua vez, a percepção musical e o processamento de melodias e harmonias, possivelmente devido à sua carga emotiva, encontra-se associado, também, a uma dominância do hemisfério direito [7]. Ademais, as estruturas associadas à percepção auditiva encontram-se no lobo temporal [8]. A amígdala, que pertence ao sistema límbico e que está fortemente associada ao processamento de emoções, e a área fusiforme da face, que está associada à percepção de faces, também se localizam no lobo temporal [8]. Consequentemente, foi encontrada uma diminuição do reconhecimento de expressões faciais em indivíduos que foram submetidos a uma excisão anteromedial do lobo temporal [9]. Músicos com ansiedade associada à performance musical demonstraram um fraco reconhecimento da expressão facial de alegria [10].

Tendo em conta que a percepção de faces e de emoções, bem como a percepção musical constituem tarefas nas quais o hemisfério direito tem uma maior dominância, com grande foco no lobo temporal, poderá ser interessante estudar de que forma, perante estímulos de expressões faciais com carga emocional, a atenção visual de indivíduos com formação

musical apresenta diferenças relativamente à atenção visual de indivíduos sem este tipo de formação.

A descodificação de expressões faciais é um processo complexo, que se inicia com a observação de vários elementos da face. A informação recolhida através do sistema visual humano é integrada e processada de modo a que o observador possa retirar uma conclusão em relação ao tipo de emoção que lhe foi apresentada e a possa classificar[4].

A acuidade visual não é muito elevada em toda a retina. Apenas na fóvea, região central da retina, existe uma resolução elevada [11]. Deste modo, durante a observação de uma imagem, apenas uma pequena região da imagem se encontrará perfeitamente definida em cada fixação do olhar do observador. Para que este consiga recolher informação suficiente para descodificar a expressão facial emocional que está a ser observada, terá que existir um varrimento da imagem com o olhar. No entanto, nem todas as imagens contêm a mesma quantidade e qualidade de informação, em particular quando estão a ser analisadas expressões faciais com diferentes cargas emocionais.

Certas regiões da face - como é o caso dos olhos e da boca- fornecem-nos mais informações do que outras e, deste modo, são consideradas mais importantes para o processo de descodificação. Porém, a atenção visual dada a cada elemento da face pode variar não só consoante o tipo de emoção representada, mas também relativamente a certas características do observador, que neste caso é o facto deste ter tido ou não formação musical [4].

Os dispositivos de *eye tracking* são ferramentas muito úteis para o estudo da atenção visual de um indivíduo perante uma imagem estímulo, pois dão-nos informações sobre a localização da atenção visual e a respetiva dinâmica. Os sistemas de *eye tracking* atualmente comercializados têm como condicionante o elevado custo associado ao equipamento, sendo que os preços destes dispositivos variam entre as centenas e os milhares de euros. O *eye tracker* utilizado neste estudo é um dispositivo de baixo custo, que foi desenvolvido, otimizado e validado no âmbito de duas dissertações para a obtenção de Grau de Mestre em Engenharia Biomédica, na FCT-UNL pelos então mestrandos Filipa Gamas [1] e Raul Vences [2].

Em 2011, foi realizado um estudo [3] cujo objetivo foi examinar a evolução da direção do olhar de indivíduos saudáveis enquanto estes examinam diferentes expressões faciais representando emoções básicas e investigar se existem padrões, na dinâmica de atenção visual, específicos a cada emoção apresentada. Os resultados deste estudo confirmam a relevância dos olhos e da boca para a descodificação das expressões, mas também demonstram que nem todas as expressões faciais com um conteúdo emocional diferente são descodificadas de igual modo. Segundo este estudo, os indivíduos olham para as regiões que são mais características de cada emoção.

A presente dissertação constitui o primeiro trabalho desenvolvido utilizando este *eye tracker* de baixo custo em contexto real. O estudo acima referido será utilizado como fio condutor para validação do dispositivo, através da comparação do resultados obtidos com os reportados em [3]. Após validada a metodologia apresentada nesta dissertação, pretende-se estudar também de que forma a formação musical de um indivíduo afeta a sua descodificação de expressões faciais.

Sendo este o primeiro trabalho levado a cabo utilizando o dispositivo em causa em contexto real, surgiram alguns problemas inesperados na recolha dos dados que tiveram de ser resolvidos, ao nível da metodologia, para que os dados fossem suscetíveis de ser utilizados e analisados, e que serão descritos no Capítulo 4-Materiais e Métodos.

1.2 Objetivos

Com este trabalho, pretende-se, perante uma amostra de dois tipos de sujeitos - com e sem formação musical -, estudar o comportamento diferenciado da atenção visual daquelas duas categorias de indivíduos, perante imagens de expressões faciais com diferentes cargas emocionais.

Para tanto, e em primeiro lugar, foi necessário validar o *eye tracker* desenvolvido em teses anteriores. Esta validação consistiu em seguir a metodologia de um estudo feito com um *eye tracker* monocular de alta velocidade e verificar que os resultados obtidos têm uma tendência semelhante aos do estudo em causa, permitindo assim concluir que os dados recolhidos pelo *eye tracker* usado neste estudo são válidos. Após esta validação do dispositivo em contexto real, é possível estudar, dentro da amostra em causa, o comportamento diferenciado de indivíduos com e sem formação musical.

CONCEITOS TEÓRICOS

2.1 O cérebro humano

O cérebro é o órgão principal e o centro do sistema nervoso do ser humano. É um órgão complexo, contém dezenas de milhares de milhões de neurónios conectados entre si, dividindo-se em dois hemisférios: o hemisfério esquerdo e o hemisfério direito. De um modo geral, o hemisfério esquerdo recebe informação e controla o lado direito do corpo e o hemisfério direito recebe informação e controla o lado esquerdo do corpo. Estes hemisférios comunicam entre si através de uma estrutura localizada na fissura longitudinal cerebral, denominada "corpo caloso"[12]. Podemos observar estas estruturas na Figura 2.1:

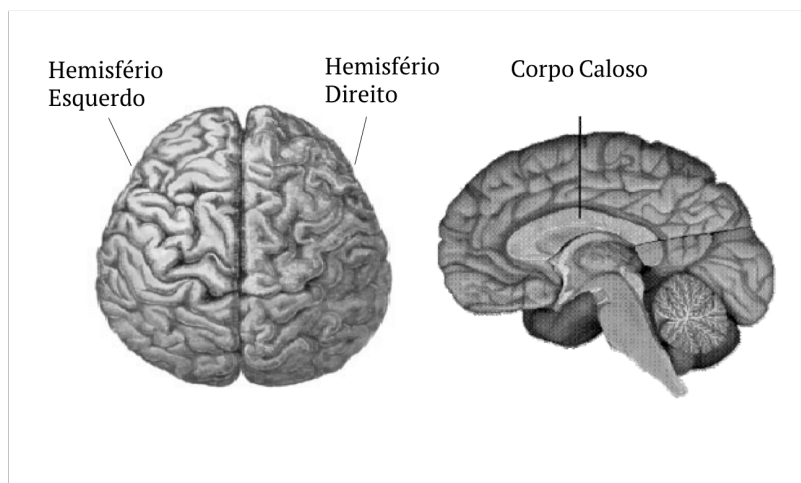


Figura 2.1: Hemisférios cerebrais e corpo caloso (adapt. [13]).

Embora a macro-estrutura dos dois hemisférios tenha uma aparência idêntica, a composição das redes neuronais de ambos é diferente, nomeadamente em termos da intensidade das ligações entre os neurónios. Esta diferenciação permite a existência de funções e processos cognitivos especializados em cada hemisfério. A este fenómeno, dá-se o nome de **lateralização da função cerebral**. Por exemplo, no que concerne a comunicação entre seres humanos, os centros de linguagem estão localizados no hemisfério esquerdo [14] enquanto que o hemisfério direito é especializado no processamento de emoções e na comunicação não-verbal [15].

Ainda que cada hemisfério seja dominante em determinadas funções, ambos têm um papel a desempenhar no processamento da informação. Acresce que o grau de lateralização cerebral não é igual para todos os indivíduos. Aqueles que têm o lado esquerdo dominante, têm tendência a ser mais metódicos e analíticos, enquanto que indivíduos cujo lado direito do cérebro é dominante tendem a ser mais criativos e artísticos [12].

Cada hemisfério está dividido em diferentes estruturas, cada uma com funções específicas. Uma destas estruturas é o lobo temporal. O lobo temporal está localizado na parte lateral do cérebro, na zona por cima das orelhas, tendo como principal função processar estímulos auditivos [16]. A amígdala, situada também no lobo temporal, faz parte do sistema límbico. Esta estrutura desempenha um papel importante relativamente às respostas emocionais de um indivíduo[8]. Ainda no lobo temporal, existe uma estrutura denominada área fusiforme da face, localizada no giro fusiforme, e tem como função a perceção de faces[17]. Na figura 2.2, podemos observar as localização das estruturas acima referidas em ambos os hemisférios.

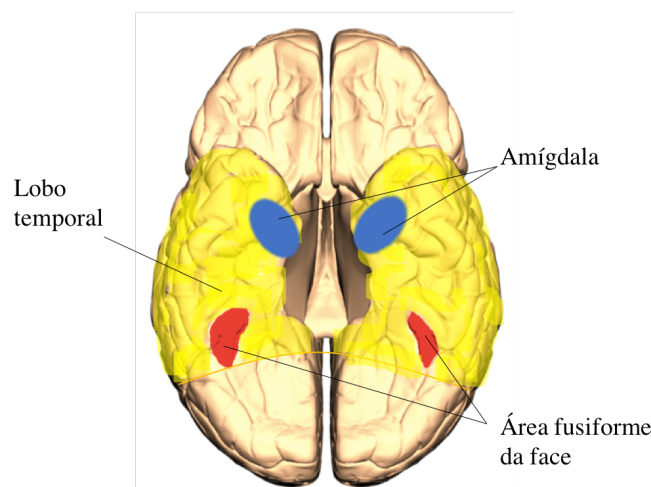


Figura 2.2: Vista ventral do lobo temporal: amígdala e área fusiforme da face (adapt. [16]).

2.1.1 Percepção e formação musical

O cérebro de cada ser humano desenvolve-se de uma forma diferente, circunstância da qual resulta uma lateralização única em cada indivíduo [12]. Esta lateralização é resultante de uma definição e discriminação das diferentes estruturas do cérebro, bem como dentro de cada uma destas regiões, devido à plasticidade cortical. A plasticidade cortical pode ser definida como a capacidade de adaptação a nível estrutural e funcional do cérebro ao longo do desenvolvimento neuronal e quando sujeito a novas experiências [18].

Ao ouvir música, o cérebro humano deteta vários elementos simples, como a afinação, acordes, ritmo e timbre. Estas características são extraídas no córtex auditivo, situado no lobo temporal, e, posteriormente, utilizadas para reconhecer a melodia e a harmonia num processo que se denomina "percepção musical". A esta percepção musical está associada uma dominância do hemisfério direito [7].

A formação musical envolve a aprendizagem de capacidades motoras e auditivas complexas, bem como o respetivo aperfeiçoamento. Indivíduos que dominam um instrumento musical, necessitam de integrar por vezes até mais do que uma linha melódica diferente ao mesmo tempo, utilizando ambas as mãos para o efeito. Para além disso, têm que, simultaneamente, ler símbolos musicais (processamento de linguagem no hemisfério esquerdo) e integrar a melodia que está escrita na sua própria interpretação (processo associado ao hemisfério direito)[19].

Vários estudos científicos demonstram que os cérebros de músicos são diferentes em relação àqueles que não possuem aptidões musicais, pois ocorrem adaptações estruturais do cérebro em resposta às capacidades trabalhadas no longo prazo [19–21]. Segundo o princípio de Hebb, a estimulação repetida e persistente da célula pós-sináptica, resulta num aumento da eficiência sináptica, ou seja, estas adaptações estruturais do cérebro decorrem da plasticidade do mesmo e desta estimulação repetida e persistente [22].

2.1.2 Descodificação de expressões faciais emocionais

As expressões faciais contêm informação relevante a nível do comportamento social dos seres humanos. Deste modo, a sua correta decodificação é importante para uma compreensão do estado emocional de quem nos rodeia. Para que as expressões faciais emocionais sejam corretamente classificadas, é necessária a integração de informação complexa. O processamento da informação facial ocorre de uma forma rápida e automática e é realizado por regiões cerebrais especializadas para o efeito [4], tendo a área fosiforme da face um papel muito importante neste processo- Figura 2.2.

O processo de reconhecimento de faces é tão espontâneo, que olhando para a Figura 2.3, podemos facilmente imaginar uma face. O ser humano tem tendência a identificar faces até em objetos inanimados quando estes apresentam algum tipo de forma comprida na horizontal (boca) abaixo de duas formas circulares lado a lado (olhos). Este reconhecimento automático deve-se a uma questão de sobrevivência bastante primária, com o objetivo de identificar o possível inimigo através do reconhecimento da face e avaliar as suas intenções através da descodificação da mesma [23].



Figura 2.3: Formato de um objeto inanimado que se assemelha a uma face.

A informação emocional extraída através da observação de uma expressão facial emocional é processada no sistema límbico, mais especificamente na amígdala [24]. O sistema límbico é composto por um conjunto de estruturas cerebrais localizadas em ambos os lados do tálamo, imediatamente abaixo do lobo temporal medial do cérebro. Este sistema desempenha diversas funções incluindo: emoção, comportamento, motivação, memória a longo-prazo e olfato. Em particular a amígdala, é uma região importante do cérebro, estando envolvida em processos de atenção e em processos emocionais [25]. Em termos cognitivos, a atenção é a capacidade que um indivíduo tem de se focar em determinados estímulos enquanto ignora outros.

A localização da atenção visual de um indivíduo perante uma expressão facial emocional é um processo espontâneo. Certas zonas da face, como é o caso da boca e dos olhos, fornecem-nos informação mais relevante e diferenciadora para uma correta interpretação da expressão facial [4]. A amígdala tem um papel importante na localização espontânea desta atenção [26].

O processo musical envolve uma elevada componente emocional e a sensibilidade musical vai sendo apurada ao longo do tempo [19]. Sendo que a amígdala tem um papel tão importante no processamento de emoções e, por sua vez, na descodificação de expressões faciais emocionais, existe uma sobreposição das regiões cerebrais envolvidas nestes

processos.

Pacientes que sofram de esquizofrenia, autismo, psicopatia ou até depressão têm um desempenho significativamente pior na classificação de expressões faciais em relação a indivíduos saudáveis. Isto deve-se ao facto de estes pacientes processarem os estímulos de uma forma diferente. Uma das diferenças neste processamento é a forma como as faces são examinadas. Enquanto que indivíduos saudáveis examinam principalmente os olhos e a boca, o traçado de olhares dos pacientes acima referidos demonstra a ausência de preferência por qualquer área específica da face [27–29].

2.2 *Eye tracking*

2.2.1 Movimentos oculares

Os olhos, além de possibilitarem a análise do ambiente envolvente, permitem discriminar os objetos quanto às suas formas, à sua distância em relação ao observador, ao facto de estarem ou não em movimento e à sua cor. Além da construção visual do ambiente onde se encontram, as imagens são utilizadas como elementos de comunicação.

O olho humano capta estímulos visuais sob a forma de luz, que é absorvida e convertida em impulsos eletroquímicos processados no cérebro- Figura 2.4. Essa luz é refratada pela córnea, estrutura avascular e transparente que funciona como um meio dióptrico. A luz entra no olho através da pupila, estrutura cujo diâmetro é variável, e a quantidade de luz que passa para a retina é controlada pela íris. Posteriormente, a luz atravessa o cristalino, que foca os objetos que se encontram a variadas distâncias. Esta luz é por sua vez projetada na retina, que tem na sua constituição fotorreceptores, os bastonetes e cones. Os bastonetes existem em maior número e reconhecem a quantidade de luz incidente, tendo pouca diferenciação no comprimento de onda. Os cones são células que têm a capacidade de diferenciar as cores e detalhes dos objetos e requerem luz relativamente brilhante para funcionar. Estas células convertem a energia luminosa em impulsos elétricos, os quais são transmitidos para o cérebro através dos neurónios da retina e do nervo ótico [11].

A fóvea é a região da retina onde existe uma concentração de cones e, devido ao facto destes serem mais sensíveis à cor e aos detalhes dos objetos, a informação é recolhida com maior resolução. Esta região da retina permite processar informação durante uma fixação. A região da retina que circunda a fóvea denomina-se parafóvea e fornece uma imagem com menor nitidez, pois existe uma maior quantidade de bastonetes e cones. À informação obtida através da parafóvea dá-se o nome de visão periférica. Apesar de a informação captada por esta região ser muito pouco detalhada, a mesma revela-se útil, uma vez que permite uma maior sensibilidade a movimentos e contrastes [31]

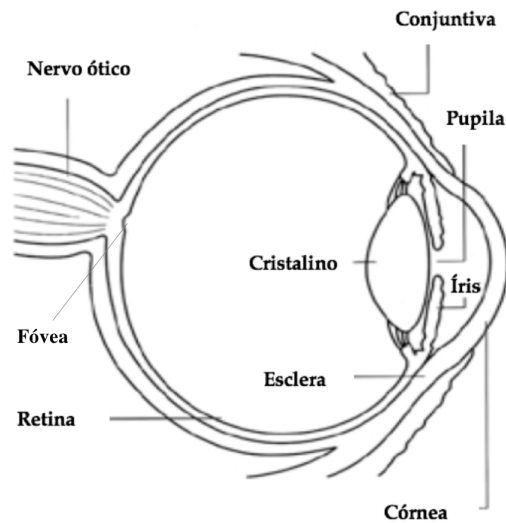


Figura 2.4: Estruturas do olho humano (adapt[30]).

De referir ainda que o processo visual se comporta de forma cíclica. Fornecido um estímulo visual, como uma imagem, toda a cena é observada, através da visão periférica, com baixa resolução. Certas regiões do campo de visão irão despertar maior interesse de um indivíduo, sendo a sua atenção direcionada para um determinado local, para este poder ser observado com maior detalhe. Haverá, assim, uma movimentação rápida dos olhos para a região que atraiu a atenção. Ao completar esse movimento, a fóvea estará direcionada para a região de interesse, de modo a que a informação possa ser processada com maior detalhe. Este processo é denominado de **fixação**, correspondendo a aproximadamente 90% do tempo de visão. Existe outro tipo de movimento, chamado **movimento sacádico**, que corresponde ao movimento rápido que o olho realiza entre fixações, de modo a reposicionar a fóvea quando a atenção visual é direcionada para outra área [32].

2.2.2 Ângulo Visual

Durante uma fixação, o olho realiza movimentos de muito baixa amplitude e elevada frequência para prevenir que objetos estacionários se desvançam. O ângulo visual é uma medida que permite apresentar a variação da posição do olho, sem que esta dependa da distância entre o estímulo e o olho- Figura 2.5.

Através do ângulo visual é possível depois calcular o tamanho da projeção para múltiplas distâncias, usando trigonometria simples (2.1): S é a região da imagem visualmente abrangida e dependendo esta da distância (D) entre o olho e a imagem e do ângulo visual (α).

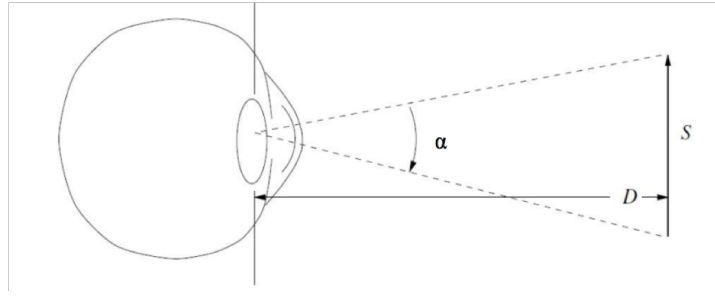


Figura 2.5: Representação do ângulo visual (adapt. [33])

$$S = 2D \times \tan\left(\frac{\alpha}{2}\right) \quad (2.1)$$

É importante notar que, em cada fixação visual, a região captada pela fóvea representa menos de 8% do campo visual.

Na Tabela 2.1 [34] estão apresentados os movimentos oculares aos quais o sistema de eyetracking utilizado na tese é sensível, destacando-se dois deles: os movimentos sacádicos e as fixações. Os **movimentos sacádicos** são movimentos rápidos e breves, atingindo velocidades de pico de centenas de graus por segundo e têm a duração de apenas dezenas de milissegundos. As **fixações** estão relacionadas com os momentos em que o olho se encontra aproximadamente imóvel, em torno de um dado alvo. Contudo, mesmo durante essas fixações, o olho realiza movimentos de muito baixa amplitude e frequência elevada. As taxas de movimento relativas à fixação e à sacada estão apresentadas na tabela em graus por milissegundos, ou seja, sob a forma de velocidade angular. A latência relativa à sacada corresponde ao tempo que decorre entre o estímulo (por exemplo, o aparecimento de um ponto no espaço) e o reposicionamento da fóvea, através do movimento sacádico, direcionada a esse estímulo.

2.2.3 Sistema de *eye tracking*

O sistema de *eye tracking* é um sistema que identifica a posição na qual o utilizador fixa o olhar e a sequência de movimentos oculares que desenvolve perante um determinado estímulo visual. Apesar de haver vários tipos de *eye trackers* não intrusivos, estes incluem dois componentes comuns: uma fonte luminosa, que é direcionada para o olho, e uma câmara, que grava a reflexão da luz juntamente com as características oculares visíveis (como a pupila) [32]

São recolhidos três parâmetros principais em cada fixação: a **localização** espacial do pixel na imagem estímulo (correspondente ao centro do foco de atenção visual), a

Tabela 2.1: Classificação dos movimentos oculares (adapt [34]).

Tipo de movimento ocular	Taxa de movimento	Latência/duração da ocorrência	Função/Significância
Fixação	<15-100 °/ms	180-275 ms	Adquirir informação, Processamento cognitivo, atenção
Sacada	<100-700 °/ms	Latência-200 ms Duração 20-200 ms	Movimento entre alvos, Reposicionamento da fóvea
Sequência de fixações	-	-	Procura direta, scanning
Tempo da direção do olhar	-	-	Processamento cognitivo
Alteração no tamanho da pupila	4,7 mm/s	140 ms	Esforço cognitivo, Representação de micro-emoções

duração dessa fixação e a **ordem temporal** pela qual as fixações aparecem, na dinâmica de varrimento visual. Através destes parâmetros podemos estabelecer várias métricas como [35]:

- Duração do olhar dentro de uma área de interesse;
- "Scanpaths"(sequência de fixações). Descreve uma sequência completa de sacada-fixação-sacada, que se traduz na transição entre áreas de interesse;
- Número total de fixações . Quanto maior for este número, menor é a eficiência da procura;
- Número de fixações sobre uma área de interesse. Tanto maior quanto maior for a respetiva importância para o utilizador;
- Duração do olhar fixo sobre uma área de interesse. Fixações longas são geralmente consideradas como indicadores da dificuldade de um participante na interpretação do conteúdo da área em causa;
- Tempo transcorrido até a primeira fixação. Quanto menor for este tempo, maior será a capacidade de as propriedades gráficas da área atraírem a atenção visual.

Sendo estes os indicadores mais populares, alguns serão mais apropriados a ser aplicados em determinadas situações, não sendo porventura tão aplicáveis em outras.

Com a informação resultante dos referidos indicadores, podemos fazer mapas com a dinâmica da exploração visual, da duração das fixações, ou ainda da informação visual que foi realmente recolhida. Estas representações gráficas demonstram, através de uma

imagem, de que modo a interface foi explorada. De seguida são apresentadas as formas das representações gráficas existentes [35]:

Traçado de olhares (*gaze plot*): é uma representação potencialmente animada, onde o ponto na interface indica o local onde o indivíduo fixou a sua atenção, em cada momento, e as linhas indicam os movimentos sacádicos- Figura 2.6.

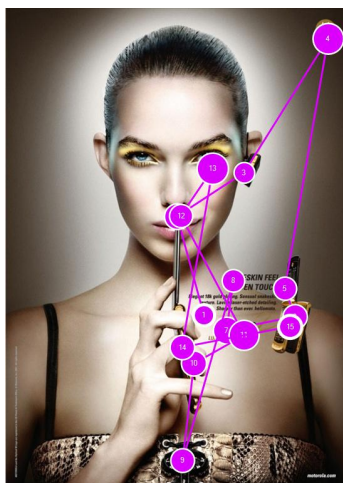


Figura 2.6: Relatório visual: traçado de olhares (*gaze plot*) (adapt. [36]).

Mapas de Calor (*heatmaps*): nestas representações, os locais onde o indivíduo fixou a sua atenção com maior frequência (áreas com elementos mais atrativos) são sinalizadas como áreas quentes ou de maior intensidade- Figura 2.7.

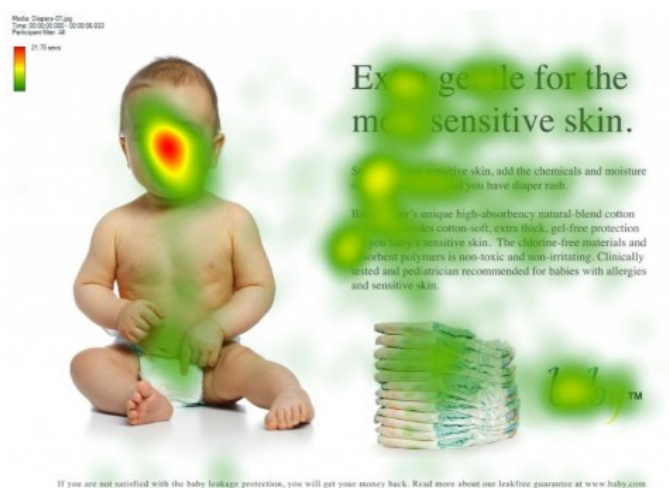


Figura 2.7: Relatório visual: *heatmap* (adapt. [36]).

Mapas de zonas sombreadas: é uma representação complementar à dos "mapas de calor" que permite visualizar em detalhe as áreas com maior concentração de fixações visuais e as zonas periféricas representadas a sombreado -Figura 2.8.



Figura 2.8: Relatório visual: mapa de zonas sombreadas (adapt. [37]).

Cluster: neste tipo de relatório são reveladas as áreas com maior concentração de pontos de fixação durante uma sessão, distribuídas percentualmente pelos participantes que demonstram interesse nas mesmas. As zonas coloridas representam os diferentes *clusters* existentes na imagem- Figura 2.9.

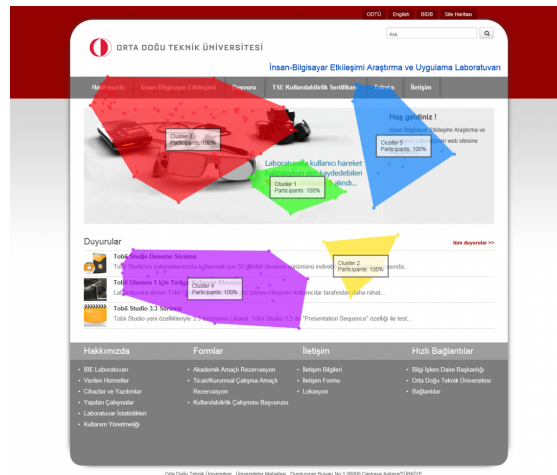


Figura 2.9: Relatório visual: *cluster* (adapt. [38]).

ESTADO DA ARTE

3.1 *Eye tracking*

Os primeiros estudos sobre o movimento ocular foram realizados por observação direta nos finais do século XIX. Louis Émile Javal, oftalmologista francês, foi o primeiro a descrever os movimentos oculares durante o ato de leitura, concluindo não se tratar de um movimento fluido, mas antes de um movimento constituído por pausas (fixações) e movimentos sacádicos [39].

Os métodos iniciais de *eye tracking* eram invasivos, envolvendo o contacto direto com a córnea. Dodge e Cline desenvolveram, em 1901, a primeira técnica *eye tracking* precisa e não-invasiva, aplicando luz refletida na córnea [40]. Em 1930, são desenvolvidos os primeiros equipamentos *eye tracking* com lentes de contacto [35].

Atualmente, os sistemas capazes de medir o movimento ocular encontram-se agrupados em 3 tipos, segundo as suas características:

Sistemas mecânicos: aplicação de dispositivos semelhantes a uma lente de contacto especial, com um sensor magnético. Trata-se de um método regularmente utilizado em investigação laboratorial relacionada com a fisiologia do olho, dado que fornece registos significativos do movimento ocular. Tem porém a desvantagem de ser muito desconfortável para os participantes durante a realização do teste- Figura 3.1.

Sistemas eletrónicos: baseia-se na utilização de potenciais elétricos, medidos a partir de elétrodos de contacto, colocados perto do olho. A variante mais comum é o eletro-oculograma (EOG) baseado no potencial elétrico permanente do olho, em que a córnea é positiva em relação à retina. Os sistemas EOG são tipicamente utilizados nos estudos

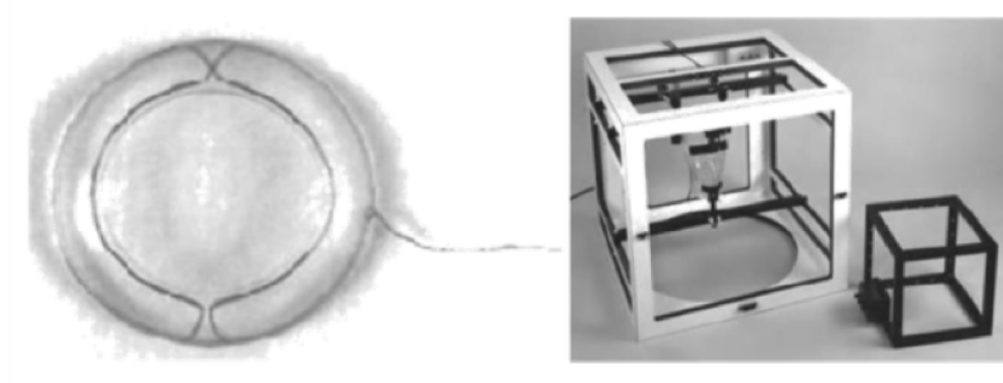


Figura 3.1: À esquerda temos a lente de contacto com o rolo magnético encapsulado e à direita o aparelho para medição dos movimentos através do campo magnético (adapt. [41]).

relacionados com o sono, pois são os mais vocacionados para medir movimentos oculares rápidos e involuntários- Figura 3.2.

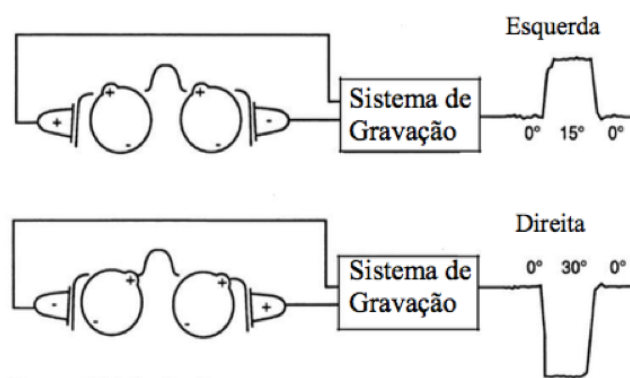


Figura 3.2: Exemplos de sinais medidos através de EOG (adapt. [42]).

Sistemas de vídeo: este tipo de equipamento é utilizado nos estudos de observação do olhar fixo e tem como vantagem o facto de não ser intrusivo. Neste tipo de sistema, a informação é analisada a partir das alterações registadas nas reflexões oculares, com base na projecção de luz no olho e na captação dos respectivos movimentos por meio de uma câmara de vídeo.

Ainda dentro dos sistemas de vídeo, existe um outro método utilizado para registar o movimento dos olhos. Este método traduz-se simplesmente em fazer o seguimento da pupila, compreendendo duas técnicas: técnica de pupila clara e técnica de pupila escura. Estas técnicas diferem na posição da fonte de iluminação relativamente ao olho, utilizando o facto da retina ser altamente reflexiva e qualquer luz que incide sobre ela é refletida de volta à fonte. A pupila aparecerá clara se a iluminação estiver alinhada com o percurso ótico e aparecerá escura se a fonte iluminante não estiver alinhada com o

percurso ótico [43].

Importa ainda destacar que os dispositivos *eye tracker* já mencionados podem ser de dois tipos: remotos ou portáteis. Sendo os primeiros totalmente não intrusivos, registando o movimento ocular à distância (normalmente colocados e integrados no computador), já os segundos requerem que o participante transporte um dispositivo próprio (por exemplo, óculos) [44]. Os dispositivos remotos geram menos ruído nos processos de avaliação, devido ao facto de serem menos intrusivos. Porém, os dispositivos portáteis têm a grande vantagem de o participante ter liberdade para se mover, podendo ser usados com menos restrições experimentais. Estes dispositivos permitem também a investigação do movimento dos olhos fora do ambiente laboratorial, bem como a investigação do papel dos movimentos oculares no dia a dia.

Hartridge e Thompson desenvolveram, em 1948, o primeiro sistema em formato de óculos, trazendo uma maior liberdade durante as experiências com *eye trackers* [45].

3.1.1 Eye tracker de baixo custo

Conforme foi já referido, foi desenvolvido um *eye tracker* de baixo custo no âmbito de uma dissertação para a obtenção de Grau de Mestre em Engenharia Biomédica, na FCT-UNL por parte da mestranda Filipa Gamas, no ano de 2017 [1]. O objetivo principal da tese em causa foi desenvolver um sistema de *eye tracking* de baixo custo, cujos componentes se resumissem a: (i) um telemóvel, o qual providencia os dados de gravação do olho, para a análise do seu movimento; (ii) um suporte em cartão, que permite manter o telemóvel numa posição fixa em relação aos olhos; e (iii) um computador, destinado a proceder à posterior análise dos dados recolhidos. Este será o sistema utilizado também no presente estudo, após algumas adaptações.

Os sistemas de *eye tracking* mais acessíveis, considerados de baixo custo, têm um valor associado na ordem das centenas de euros. Ainda que esta gama de preços seja a menos dispendiosa disponível no mercado, continua a ser um preço globalmente elevado. O sistema proposto nessa dissertação, admitindo que o utilizador deste sistema já possui um telemóvel e um computador compatíveis, tem apenas o custo relativo ao suporte em papel cartonado, o qual corresponde a menos de 10€.

Este *eye tracker* de baixo custo foi desenvolvido para funcionar na seguinte forma:

1. O telemóvel é colocado dentro do suporte de cartão-Figura 3.3.
2. São apresentados ao utilizador nove pontos de calibração no ecrã do telemóvel. Para cada ponto de calibração, é recolhida uma fotografia do olho do utilizador, com recurso à câmara frontal do telemóvel-Figura 3.4.

3. Após captação das nove imagens de calibração, a imagem estímulo é apresentada no ecrã do telemóvel. Enquanto esta imagem está a ser apresentada, os movimentos oculares são filmados com a câmara frontal do telemóvel- Figura 3.5.
4. São carregados para a interface os vários ficheiros: nove imagens de calibração, o vídeo com os movimentos oculares e estímulo.
5. Através do algoritmo de análise, é encontrado o centro da pupila para as nove imagens de calibração. Esta informação serve para encontrar a função que faz corresponder qualquer ponto do centro da pupila com a localização da direção do olhar no estímulo. De seguida, encontra-se o centro da pupila para todas as *frames* do vídeo que registou os movimentos oculares.
6. Finalmente, tendo disponíveis todas as posições da pupila e a função de transformação geométrica é possível mostrar a direção do olhar no estímulo ao longo do tempo, percebendo assim quais as zonas de maior interesse no referido estímulo [1].



Figura 3.3: Colocação do telemóvel dentro do suporte.

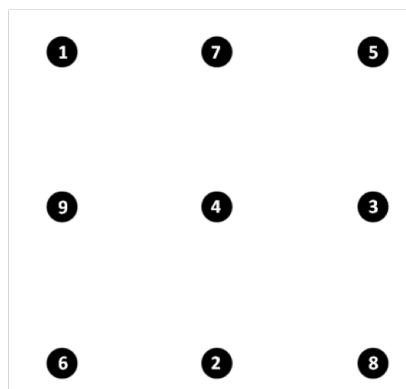


Figura 3.4: Ordem de apresentação dos pontos de calibração [1].



Figura 3.5: Utilização do dispositivo durante a apresentação do estímulo.

Apesar de este sistema de *eye tracking* ter apresentado, de acordo com a dissertação acima referida, bons resultados, a própria dissertação admite existirem algumas limitações no algoritmo, que acabam por refletir na precisão e exatidão dos resultados. Ademais, o sistema em causa estava adaptado apenas ao modelo do telemóvel utilizado durante o referido trabalho, pelo que qualquer outro telemóvel com diferentes especificações, nomeadamente no que toca ao tipo de câmara frontal, não teria uma *performance* tão boa.

No ano seguinte, o mencionado *eye tracker* de baixo custo foi validado e otimizado, no âmbito da dissertação para obtenção de Grau de Mestre em Engenharia Biomédica pelo então mestrando Raul Vences [2].

O sistema sofreu alterações profundas com vista a tornar o sistema mais robusto e adaptado a um maior número de telemóveis. Após terem sido feitas algumas modificações no sistema, obteve-se um dispositivo com uma exatidão entre 1.06-1.51 graus, e uma precisão entre 0.6-0.86 graus, sendo de referir a este respeito que os sistemas comerciais apresentam, normalmente, uma exatidão entre os 0.5-1.0 graus e uma precisão inferior a 0.3 graus. Assim, estes resultados permitiram validar o sistema de *eye tracker*, apesar de este apresentar uma *performance* ligeiramente inferior à dos sistemas comerciais. Além disso, pode afirmar-se que o dispositivo testado se encontra em condições de ser utilizado em outros estudos de atenção visual.

3.2 Decodificação de expressões faciais emocionais

As expressões faciais são uma das formas mais importantes que o ser humano tem para comunicar o seu estado emocional. Na cultura popular e na literatura científica sobre o processamento de faces, a área dos olhos é frequentemente considerada a parte mais importante para o reconhecimento de expressões faciais.

Vários estudos concluíram que a localização das informações úteis nas diferentes áreas

faciais para a descodificação de expressões faciais não é igual para todas as expressões. Ou seja, algumas áreas do rosto podem transmitir mais informação do que outras na discriminação das emoções apresentadas [46–48].

Para além disto, foi demonstrado que em muitas patologias neurológicas, a região ocular é processada com menor eficiência ou de uma maneira anormal [49, 50].

Um paciente com lesão bilateral da amígdala [49] mostrou-se incapaz de descodificar expressões de medo em virtude de não fixar espontaneamente a sua atenção visual nos olhos durante a análise da expressão facial. Porém, quando lhe foi instruído fixar a sua atenção na região dos olhos, este indivíduo foi capaz de identificar a expressão facial corretamente. Este caso não só valida o facto de que a amígdala tem um papel importante nos mecanismos envolvidos na atenção dada a cada elemento da face, como a circunstância de certas áreas faciais fornecerem informações importantes para a descodificação das emoções representadas.

A partir das observações acima listadas, é tentador concluir que os olhos são mais importantes do que qualquer outra área facial para a descodificação das emoções faciais. No entanto, os estudos que abordaram diretamente a questão de saber quais as características faciais úteis para a discriminação de expressões faciais emocionais, chegaram a resultados contraditórios. Alguns concluíram que a parte inferior da face era mais importante do que a sua parte superior [51, 52], enquanto outros não encontraram uma maior importância de uma parte da face relativamente a outra [53, 54].

Estudos de *eye tracking* que examinaram de que forma as fixações oculares são distribuídas nos rostos durante o reconhecimento de expressões faciais, encontraram uma amostra aproximadamente igual das áreas da boca e dos olhos[55].

Bassili, 1979, mostrou que, para algumas emoções (alegria e nojo), as informações fornecidas pela parte inferior da face são cruciais, enquanto outras emoções (tristeza e medo) são reconhecidas de uma forma mais eficaz através da observação da face superior[56].

3.2.1 Hemisfério direito

Estudos clínicos e experimentais sugerem que o hemisfério direito está envolvido preferencialmente no processamento da emoção nos humanos [57–59]. Estudos anteriores mostraram que o dano no hemisfério direito pode prejudicar o processamento de faces ou cenas emocionais [60] e que a estimulação elétrica do cortex visual do lobo temporal direito pode dificultar o processamento de expressões faciais [61].

3.2.2 Monitorização dos movimentos oculares na decodificação de expressões faciais

Na presente dissertação, de modo validar a metodologia de eyetracking empregue, uma primeira experiência é efetuada seguindo a metodologia aplicada num estudo desenvolvido em 2011 [3]. Esse estudo, teve como objetivo examinar o traçado de olhares de indivíduos saudáveis enquanto estes observavam expressões faciais emocionais e investigar se existem traçados de olhares específicos para cada emoção.

Caso os resultados obtidos coincidam com os resultados encontrados em [3], fará sentido estender a metodologia ao estudo entre duas populações com formação artística diferenciada (neste caso, com e sem formação musical) de forma a observar se existe alguma diferença entre a forma como estas duas populações explora visualmente as imagens, bem como a que regiões é dada uma maior atenção visual.

MATERIAIS E MÉTODOS

4.1 Materiais

Os materiais envolvidos nesta dissertação são: os três componentes do sistema de *eye tracking* (um telemóvel, um suporte e um computador); as imagens estímulo, que fazem parte de um banco de imagens denominado "*Karolinska Directed Emotional Faces*", frequentemente utilizado em estudos desta natureza e duas escalas para a avaliação da valência e da ativação das imagens estímulo denominadas "*Self Assessment Manikin (SAM)*".

4.1.1 Telemóvel

De forma a recolher os dados do movimento oculomotor dos participantes, foi utilizada uma versão otimizada [2] da aplicação para Android desenvolvida pela então mestrande Filipa Gamas [1]. A recolha exige o recurso a um telemóvel que tenha a aplicação instalada e que, através da sua câmara frontal, faça o registo de vídeo de um dos olhos, enquanto o participante é submetido ao estímulo. Na Figura 4.1 podemos observar a interface da aplicação Android.

4.1.2 Suporte

O suporte consiste num *cardboard*, sendo este fixado à cabeça através de um fita, de modo a que a posição entre o ecrã e os olhos do participante não se altere com o movimento, como é possível observar na Figura 4.2. Este fator é importante para que, uma vez feita a calibração para o participante, esta se mantenha válida durante todo o tempo do estímulo.

Em relação aos métodos descritos acima para o seguimento da direção da atenção visual, a técnica utilizada foi a do seguimento da pupila, mais concretamente, o da pupila escura. O suporte dispõe de lentes que possibilitam a focagem na retina dos objetos

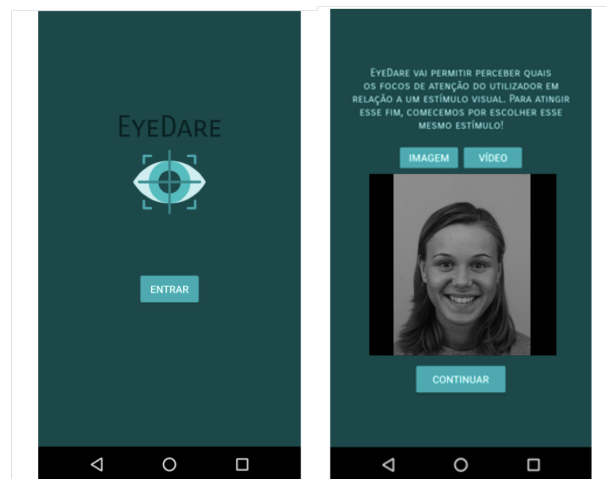


Figura 4.1: À esquerda entrada na aplicação e à direita escolha de um estímulo.

apresentados no ecrã, visto que a distância ao estímulo é inferior ao ponto próximo de visão (ponto mais próximo do olho onde ainda é possível observar um objeto com nitidez). O suporte para o telemóvel proporciona ainda um isolamento em relação ao ambiente exterior, fazendo com que as condições de iluminação dentro do suporte sejam aproximadamente constantes, independentemente do local onde o participante se encontra. Este isolamento permite, por sua vez, também que não existam distrações para o participante resultantes do ambiente envolvente. O *cardboard* tem embutido um circuito eletrónico de iluminação, alimentado por duas pilhas de 3V, que permite o funcionamento em paralelo de dois LED's de alto brilho, um para cada lente. Estes LED's, permitem ter uma imagem mais detalhada dos olhos, assegurando que ambos estejam expostos a condições de luminosidade semelhantes [2].

Para a câmara ficar mais centrada com o olho, o suporte possui uma calha sob o local onde se coloca o telemóvel. Por questões de conforto e estabilidade, o suporte possui ainda um apoio para o nariz, por forma a que a fita não exerça uma pressão excessiva sobre o nariz do utilizador, auxiliando também a manter a cabeça deste e o *cardboard* sempre na mesma posição relativa.

4.1.3 Computador

Para proceder à análise da resposta aos vídeos de calibração e estímulo foi necessário recorrer a um computador, de modo a determinar os locais que foram fixados durante a visualização de cada estímulo. Na Figura 4.3 está representada a interface utilizada que foi criada utilizando a funcionalidade GUI-*Graphical Interface User* do Matlab.



Figura 4.2: Diferentes perspetivas do suporte em cartão. À esquerda podemos ver o interruptor do sistema e o suporte das pilhas, ao centro podemos ver a fita e o suporte do nariz e à direita podemos ver o interior do *cardboard* iluminado (adap. [2]).

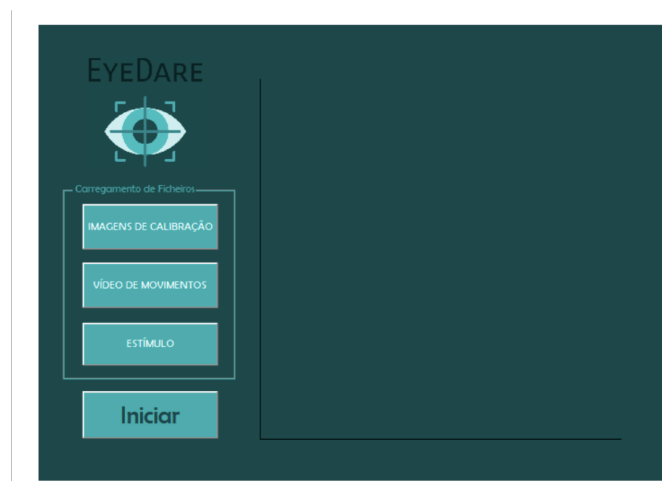


Figura 4.3: Interface através da qual é feito o carregamento do vídeo de calibração, do vídeo de movimentos correspondente ao estímulo e do próprio estímulo [2].

4.1.4 *Karolinska Directed Emotional Faces*

O *Karolinska Directed Emotional Faces* [62] é um conjunto de 4900 fotografias de expressões faciais humanas. Este conjunto de imagens contém fotografias de 70 atores amadores representando 7 emoções diferentes, sendo cada expressão apresentada através de 5 ângulos diferentes.

Nesta dissertação utilizaram-se apenas 80 dessas imagens. Foram selecionadas as imagens referentes ao ângulo frontal, onde 16 atores (8 do sexo masculino e 8 do sexo feminino) representam expressões faciais das 5 emoções básicas: Alegria, Tristeza, Neutro, Medo e Raiva. Na Figura 4.4, podemos observar alguns exemplos das imagens estímulo utilizadas.



Figura 4.4: Expressões faciais emocionais. Da esquerda para a direita: Neutro, Tristeza, Medo, Raiva e Felicidade [62].

4.1.5 *Self Assessment Manikin*

O *Self Assessment Manikin* (SAM) [63] foi o procedimento utilizado nesta dissertação para a medição da **valência** e **ativação** das imagens estímulo por parte de cada participante. A avaliação da valência é feita em relação à emoção representada na imagem estímulo, ou seja, quão positiva ou negativa a emoção expressa é. A ativação corresponde a uma excitação interna desencadeada no participante pela imagem estímulo. Este procedimento, que é apresentado numa escala de 1 a 9 pontos, consiste em dois pictogramas, um relativo à valência do estímulo, Figura 4.5, e outro relativo à sua ativação, Figura 4.6. A escala referida tem como vantagem o facto de constituir um procedimento não só quantitativo, mas também rápido e intuitivo através da observação do pictograma.

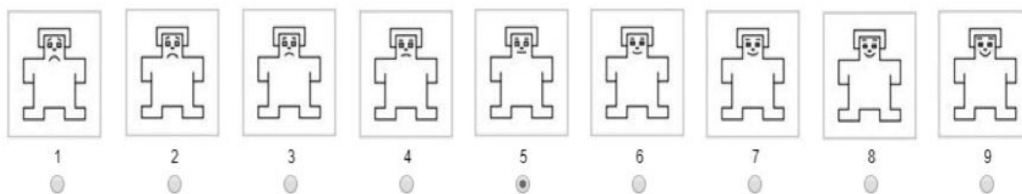


Figura 4.5: Escala utilizada para a avaliação de valência.

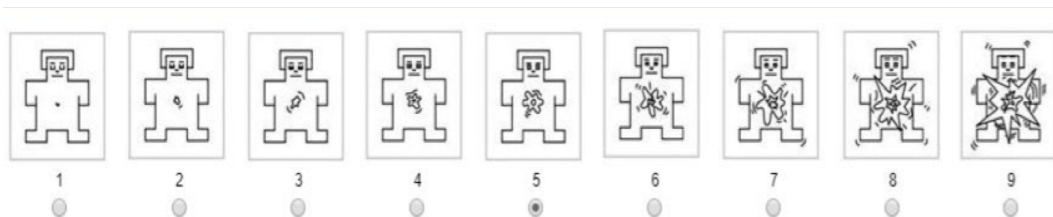


Figura 4.6: Escala utilizada para a avaliação de ativação.

4.2 Métodos

A utilização do sistema de *eye tracking* compreende duas fases, a primeira das quais consiste na recolha de dados. Os dados relativos aos movimentos oculares foram recolhidos no telemóvel através da aplicação Android já referida. A segunda fase é aquela na qual, através de um algoritmo associado a uma interface em Matlab (no computador), o seguimento da pupila é traduzido em coordenadas XY, no plano das imagens apresentadas, ao longo do tempo de exposição utilizado.

Segue-se um pré processamento dos dados obtidos. Verificou-se que a posição relativa entre o telemóvel e o olho, durante cada recolha, nem sempre se manteve inalterada. Teve de existir, portanto, alguma recalibração dos dados para que estes pudessem ser utilizados. Este processamento será explicado em maior detalhe no ponto 4.2.3.

Finalmente, aplicaram-se várias métricas, descritas no ponto 4.2.4, aplicadas aos dados obtidos para se proceder a uma análise comparativa dos resultados.

4.2.1 Participantes

Foram efetuadas recolhas com 46 participantes. Porém, 16 destas recolhas não puderam ser utilizadas, devido a descalibrações irreversíveis, resultantes de uma certa fragilidade do dispositivo. Sendo um dispositivo pouco confortável, alguns participantes sentiram a necessidade de o ajustar após a calibração, conduzindo a que os dados recolhidos não pudessem ser utilizados. Consequentemente, foram utilizadas para este estudo as recolhas feitas com 30 indivíduos voluntários.

Dos 30 participantes, 19 são do sexo masculino e 11 do sexo feminino, com idades compreendidas entre os 17 e os 30 anos ($22,70 \pm 2,89$). A amostra de participantes divide-se em dois grupos: um com formação musical e outro sem formação musical, cada grupo tem 15 participantes. Para o presente estudo, foram considerados indivíduos com formação musical aqueles que tiveram pelo menos 4 anos de aprendizagem de um instrumento musical, iniciada com uma idade inferior aos 12 anos. Dentro da amostra de participantes com formação musical, o número de anos e o grau de formação atingido apresenta uma grande variabilidade, indo desde os 4 anos de formação até à conclusão do último grau do conservatório de música.

Os participantes foram previamente informados que os seus movimentos oculares iriam ser monitorizados durante uma tarefa de observação de expressões faciais. Foi-lhes também explicado que, após cada imagem estímulo, teriam que a classificar em relação a duas escalas (valência e ativação) utilizando o SAM.

4.2.2 Recolha de dados

Após ter sido feita a explicação do estímulo, o dispositivo de *eyetracking* foi colocado, sendo iniciada a sua calibração. Posteriormente, foram apresentadas as imagens estímulo. Depois de cada imagem estímulo, o participante classificou a expressão facial em relação à sua valência e, de seguida, em relação à ativação provocada pelo estímulo. Esta avaliação foi feita oralmente, sem ter sido retirado o dispositivo. Cada imagem estímulo foi apresentada durante 2,5 segundos, demorando cada processo global de recolha de dados cerca de 14 minutos.

Os dados recolhidos pelo telemóvel foram processados em Matlab onde, através de um algoritmo desenvolvido e otimizado em teses anteriores [1, 2], se obtiveram as coordenadas da localização da atenção visual para cada *frame* do vídeo, dos movimentos oculares recolhidos.

Idealmente, os dados recolhidos através do dispositivo têm um aspeto semelhante ao da Figura 4.7.

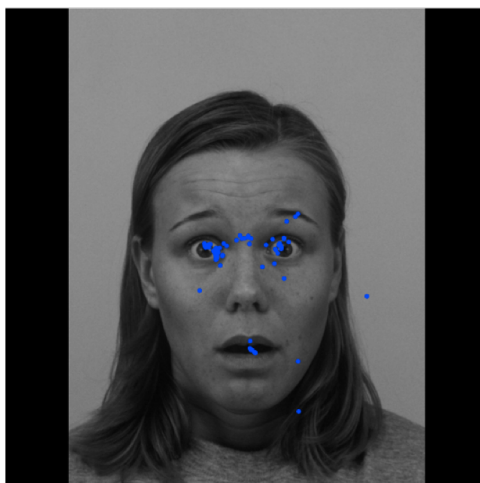


Figura 4.7: Exemplo da apresentação dos dados recolhidos em eyetracking

Porém, devido à sensibilidade do dispositivo a pequenos movimentos, algumas das recolhas apresentaram um desvio da localização do conjunto de pontos apresentados. Os conjuntos de pontos deslocados tiveram que sofrer uma translação para compensar este desvio. Esta translação será explicada em maior detalhe no ponto 4.2.3.

4.2.3 Pré-processamento dos dados

Em resultado da circunstância de ter sido o primeiro estudo realizado com este dispositivo em contexto real, verificaram-se alguns problemas associados à sua utilização que não

foram detetados nos estudos anteriores [1, 2]. Foram identificadas três situações diferentes, que resultaram num desvio da estimativa da localização das coordenadas referentes à atenção visual do participante.

Um destes problemas deriva do facto do tempo de utilização do *eye tracker* ser bastante superior ao tempo de utilização do dispositivo em testes anteriores. Como este dispositivo assenta na face durante bastante tempo, fruto da pressão exercida, a pele vai-se adaptando à forma do suporte de cartão, fazendo com que a posição deste desça em relação à posição dos olhos do participante, de uma forma aproximadamente contínua e linear. Isto resulta num desvio da estimativa da localização dos pontos obtida em relação à localização da atenção visual real. Este fenómeno não ocorreu em todas as recolhas, visto que a anatomia facial de cada participante determina a forma como o dispositivo assenta na face e, em determinados casos, não houve alteração da posição relativa entre o *cardboard* e os olhos do participante.

Para se proceder à correção deste desvio, teve de se efetuar uma translação gradual dos pontos segundo a direção vertical (yy). Como a variação da posição do *cardboard* é um processo gradual e linear, a translação necessária para corrigir a posição dos pontos obtidos deve ser também gradual e linear. Para tal, o procedimento utilizado para este caso consistiu em avaliar qual foi o desvio do conjunto de pontos correspondentes à última imagem-estímulo e construir uma reta de ajuste à estimativa vertical. O declive da reta de ajuste será o resultado da divisão do desvio considerado para a última imagem-estímulo pelo número de imagens estímulo. É retirado a cada ponto um determinado valor, sendo que a medida de translação para cada ponto é progressiva e segue uma reta de proporcionalidade direta. Na Figura 4.8, podemos observar a translação gradual dos pontos de uma recolha, em diferentes momentos da mesma.

Um outro problema detetado resulta do facto dos estímulos serem imagens com carga emocional. Assim, alguns dos participantes reagiram a estas imagens movimentando a própria face, por exemplo através de um sorriso ou um franzir das sobrancelhas. Esta movimentação dos músculos faciais provoca uma variação da posição relativa entre o suporte e os olhos do participante, não seguindo nenhum padrão em particular, dependendo apenas da reação de cada participante perante cada imagem estímulo.

A correção destes desvios teve de ser feita manualmente, pois a aceleração provocada pelos músculos faciais é muito menor do que aceleração dos movimentos oculares, resultando em erros com aparência aleatória, e inviabilizando o desenvolvimento de um algoritmo de correção em tempo útil. Na Figura 4.9, é apresentado um exemplo de uma destas correções efetuadas manualmente.

Existiu ainda um outro problema, menos frequente, que afetou apenas três recolhas.

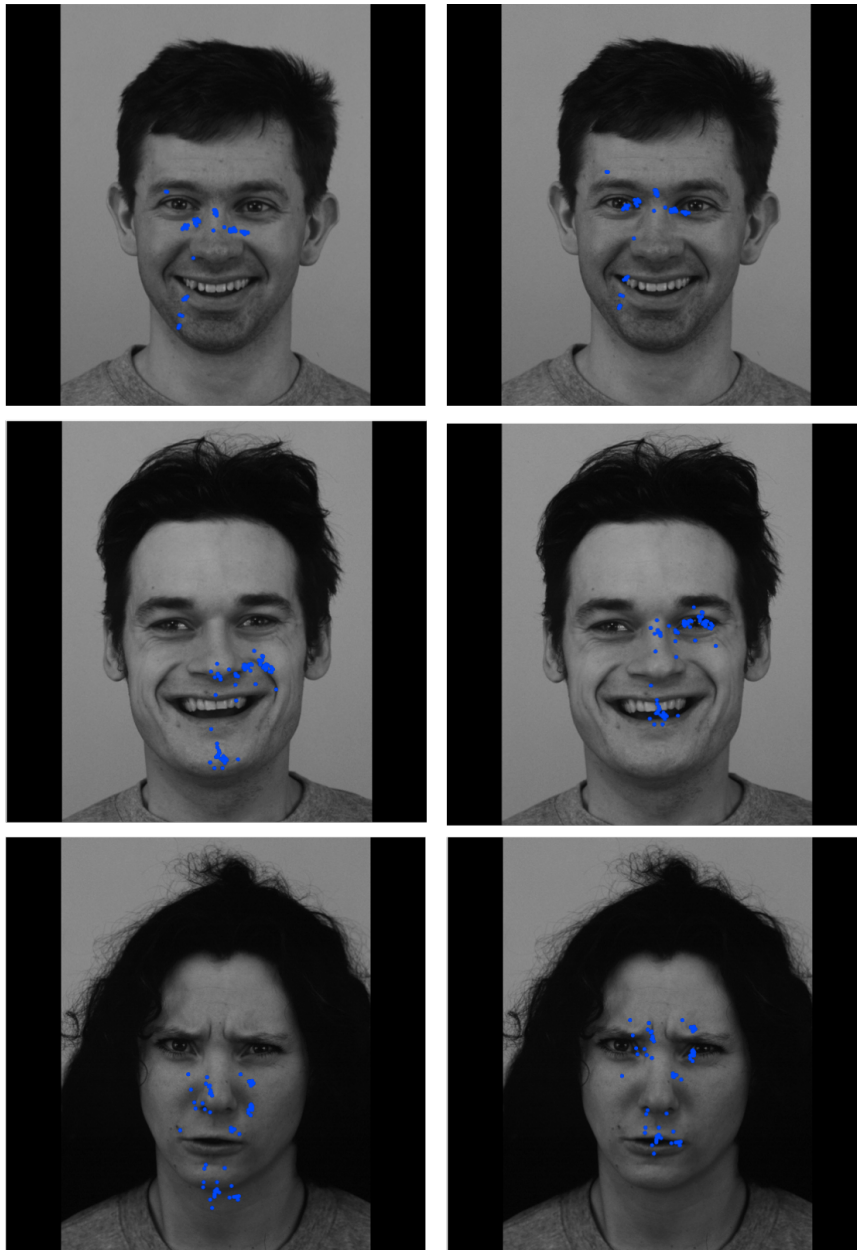


Figura 4.8: Apresentação dos dados antes (à esquerda) e após (à direita) a translação. De cima para baixo: imagem-estímulo nº 15, nº 40 e nº 80.

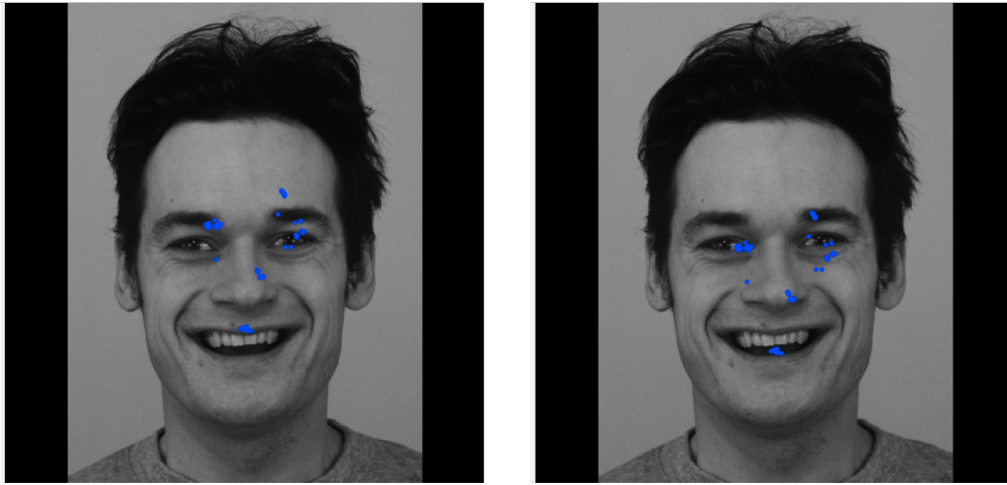


Figura 4.9: Translação manual. À esquerda antes do processamento e à direita após o processamento

Como o telemóvel assenta numa calha, dentro do suporte de cartão, qualquer deslize na calha conduz a um desvio das coordenadas em xx . Para corrigir este desvio, foi necessário identificar o momento em que este deslize ocorreu e fazer uma translação de todos os pontos a partir desse momento, para todas imagens estímulos seguintes. Na Figura 4.10 é possível observar o desvio dos dados provocado pelo deslize do telemóvel e a sua respetiva correção.

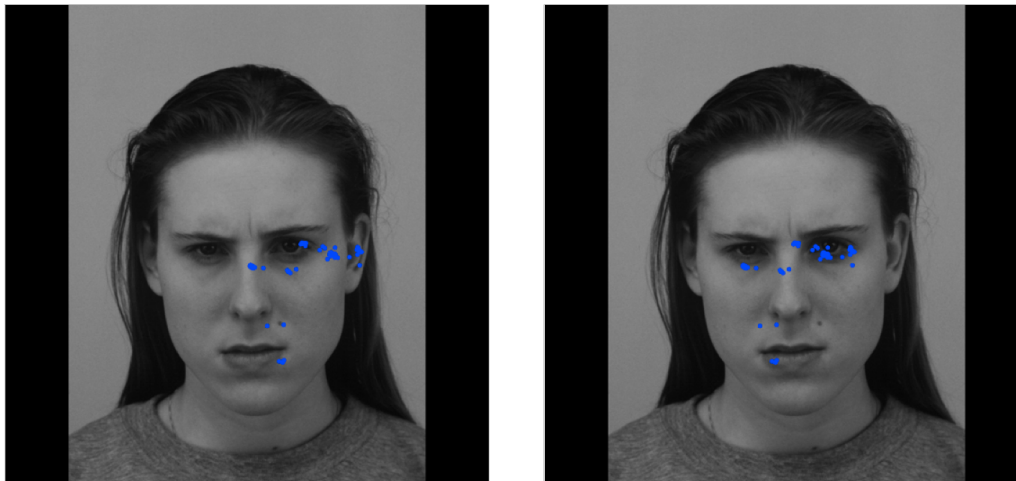


Figura 4.10: Translação horizontal. À esquerda antes do processamento e à direita após o processamento

4.2.4 Análise de dados

Para cada frame do vídeo que gravou a movimentação do olho, o algoritmo em Matlab traduz a posição da pupila em coordenadas na imagem estímulo

O sistema de *eye tracking* utilizado no presente estudo encontra a posição da pupila (traduzindo esta posição para coordenadas xy na imagem estímulo) aproximadamente a cada 0,33 segundos. Assim, nem todos os pontos exibidos na imagem-estímulo correspondem a fixações. Alguns destes pontos correspondem à movimentação do olho entre fixações (movimentos sacádicos) e não são importantes para a avaliação da atenção visual do participante, visto que estes movimentos são meramente transitórios e não têm como função recolher informação da imagem. Porém, como referido anteriormente, durante uma fixação, o olho realiza movimentos de muito baixa amplitude e elevada frequência para prevenir que objetos estacionários se desvançam. Atendendo à metodologia utilizada em [3], definiu-se como fixação um olhar que apresentasse uma variação da posição do olho não superior a 0,8 graus, correspondente a um diâmetro inferior a 18 pixels na imagem-estímulo, durante pelo menos 100 milissegundos.

Como o objetivo do estudo é perceber a relevância de certas áreas faciais na descodificação de expressões faciais, foram definidas oito áreas de interesse: olhos esquerdo, olho direito, face esquerda, face direita, testa, nariz, boca e queixo. Os limites das áreas de interesse encontram-se definidos na Figura 4.11. Estes limites foram definidos manualmente para cada imagem-estímulo, de modo a que cada região de interesse fosse corretamente delimitada.

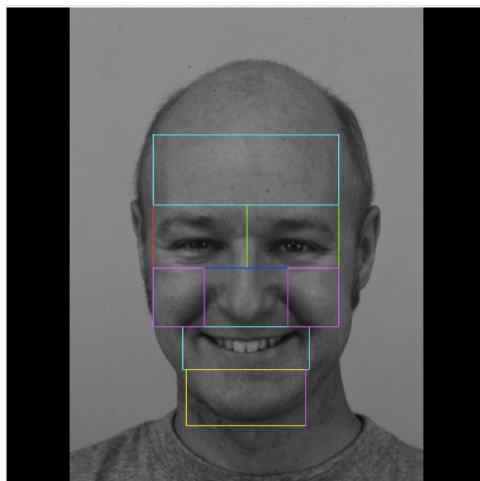


Figura 4.11: Limites das áreas faciais de interesse.

Foi estudado o efeito da emoção na localização das fixações e foi igualmente estudada a localização das fixações de um modo geral. Para tal, foram definidas as seguintes métricas:

Localização da primeira fixação: É registada a primeira fixação relativa cada imagem-estímulo de modo a comparar as frequências de fixação entre as diferentes áreas de interesse para cada tipo de emoção apresentada. A localização da primeira fixação é importante para se poder retirar conclusões acerca da zona da face que atrai mais a atenção

do participante. Sendo esta a primeira informação a ser processada na decodificação da expressão facial emocional apresentada ao participante, é uma métrica indicativa da prioridade que este demonstra ao examinar a expressão facial.

Duração média das fixações por área de interesse: Esta métrica compara o tempo durante o qual os participantes fixaram determinada zona da face em relação ao tempo total de fixação. A partir desta informação é possível perceber que zonas da face são fixadas durante mais tempo e se existe variação desta duração em função da emoção apresentada ao participante. Assim, é importante para se perceber se a emoção apresentada influencia ou não o tempo que o participante despende a examinar cada zona da face e que zonas são analisadas durante mais tempo pois captam mais a atenção do indivíduo.

Número médio de fixações por área de interesse: Esta métrica compara o número de fixações em cada zona da face em relação ao número de fixações total. A partir desta informação é possível perceber que zonas da face foram "visitadas" mais vezes comparativamente às restantes. Deste modo, serve para perceber que áreas da face são mais importantes para a decodificação de expressões faciais e se existe uma variação significativa das áreas mais fixadas em função da emoção apresentada ao participante.

Rácio médio da duração das fixações entre os olhos e a boca: Esta métrica compara a duração das fixações na boca e nos olhos, dividindo a diferença entre a duração de fixação nos olhos e a duração de fixação na boca pela soma da duração de fixação nos olhos com a duração na de fixação boca. É uma métrica importante pois os olhos e a boca têm tendência a ser as zonas mais atrativas no processo de decodificação de imagens e, através deste rácio, é possível perceber se existe uma variação da importância relativa destas áreas consoante a emoção apresentada.

Distância percorrida pelo olho: Esta métrica regista a distância percorrida pelo olhar durante a decodificação das diferentes expressões faciais. É importante para avaliar se certas emoções provocam uma movimentação dos olhos mais dispersa ou uma movimentação mais direcionada e contida.

De modo a estudar o efeito dos fatores **emoção** e **área de interesse** nos resultados obtidos, foi realizada uma análise de variância (ANOVA) e um teste-F para cada métrica, para aferir se os fatores emoção e área de interesse tiveram ou não um efeito significativo nos resultados obtidos. Para além disso, foram efetuados testes-t para comparar as diferenças entre as diferentes áreas de interesse ou as diferentes emoções. O *valor – p* obtido serviu para determinar a significância dos resultados, sendo que se considerou que o resultado do teste é estatisticamente significativo quando $p < .05$. De igual modo, relativamente à análise de variância, também foi apresentado o valor de $\eta^2_{parcial}$, sendo este a razão de variação associada a um efeito, mais esse efeito e a sua variação de erro

associada.

RESULTADOS E DISCUSSÃO

Neste capítulo são apresentados e discutidos os resultados do estudo que constituiu esta dissertação. Como já referido anteriormente, foi seguida a mesma metodologia de um estudo realizado em 2011 [3], com o objetivo de validar o dispositivo de *eye tracking* de baixo custo e também validar o procedimento utilizado no referido estudo. Assim, os resultados são apresentados e comparados com os resultados do referido estudo. Após esta análise comparativa e consequente validação do dispositivo, foi feita uma análise comparativa do comportamento diferenciado entre o grupo de participantes com formação musical e sem formação musical.

5.1 Localização da primeira fixação

Procedeu-se à construção de uma tabela de frequências da primeira fixação em cada área de interesse para as diferentes emoções. Analisou-se o fator área de interesse, o fator emoção e a interação entre os dois fatores.

Na Figura 5.1, estão apresentadas as médias da frequência da primeira fixação para cada área de interesse e para cada emoção, bem como o desvio padrão associado a cada média. No gráfico podemos observar uma predominância evidente dos olhos como alvo da primeira fixação relativamente a qualquer outra área de interesse.

Verificou-se um efeito significativo do fator área de interesse, $F(4,116) = 84.18; p < .001; \eta^2_{\text{parcial}} = .74$, e uma interação significativa entre o fator emoção e o fator área de interesse, $F(16,464) = 2.00; p = .001; \eta^2_{\text{parcial}} = .07$. Não foi encontrado um efeito significativo do fator emoção.

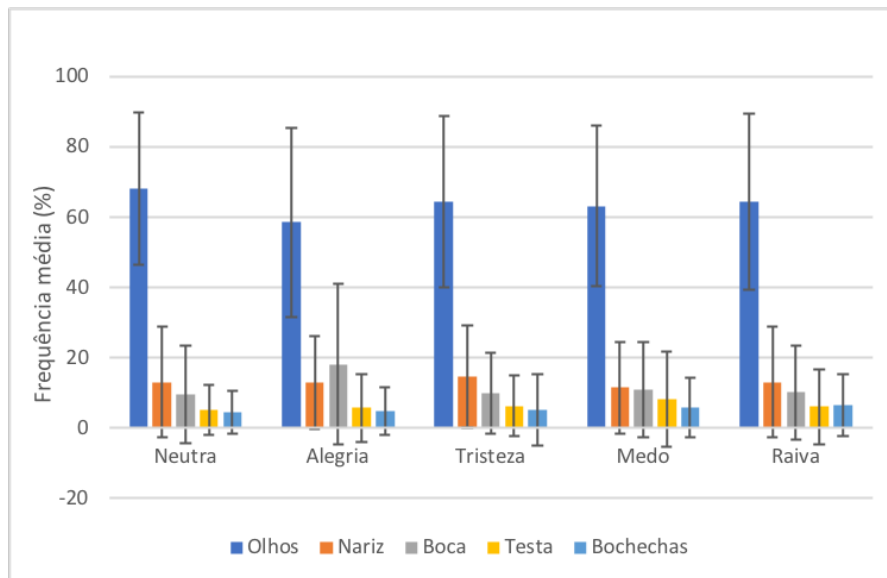


Figura 5.1: Valores médios da frequência da primeira fixação para cada área de interesse relativos às diferentes emoções.

Verificou-se que a primeira fixação é direcionada a maioritariamente aos olhos independentemente do tipo de emoção apresentada. Os olhos são portanto mais significativamente alvo da primeira fixação em relação: à testa, $T(29) = 13.11; p < .001$, às bochechas, $T(29) = 12.68; p < .001$, ao nariz, $T(29) = 8.79; p < .001$ e à boca, $T(29) = 9.02; p < .001$. Foi também encontrado um efeito significativo da boca em relação às bochechas, $T(29) = 2.86; p < .001$, mas este não se verificou em relação ao nariz e à testa.

5.1.1 Validação geral

Estes resultados seguem a mesma tendência dos resultados encontrados em [3], pois foi também encontrado um efeito significativo do fator área de interesse e da interação entre o fator área de interesse e o fator emoção. Também no referido estudo, não se verificou um efeito significativo do fator emoção. Os olhos foram também o alvo mais frequente da primeira fixação na face.

5.1.2 Efeito da formação musical

A amostra representada em 5.1 foi dividida entre dois grupos, um com formação musical o outro sem este tipo de formação. Na Figura 5.2, podemos observar as diferenças entre frequências médias da primeira fixação relativamente aos grupos com e sem formação musical.

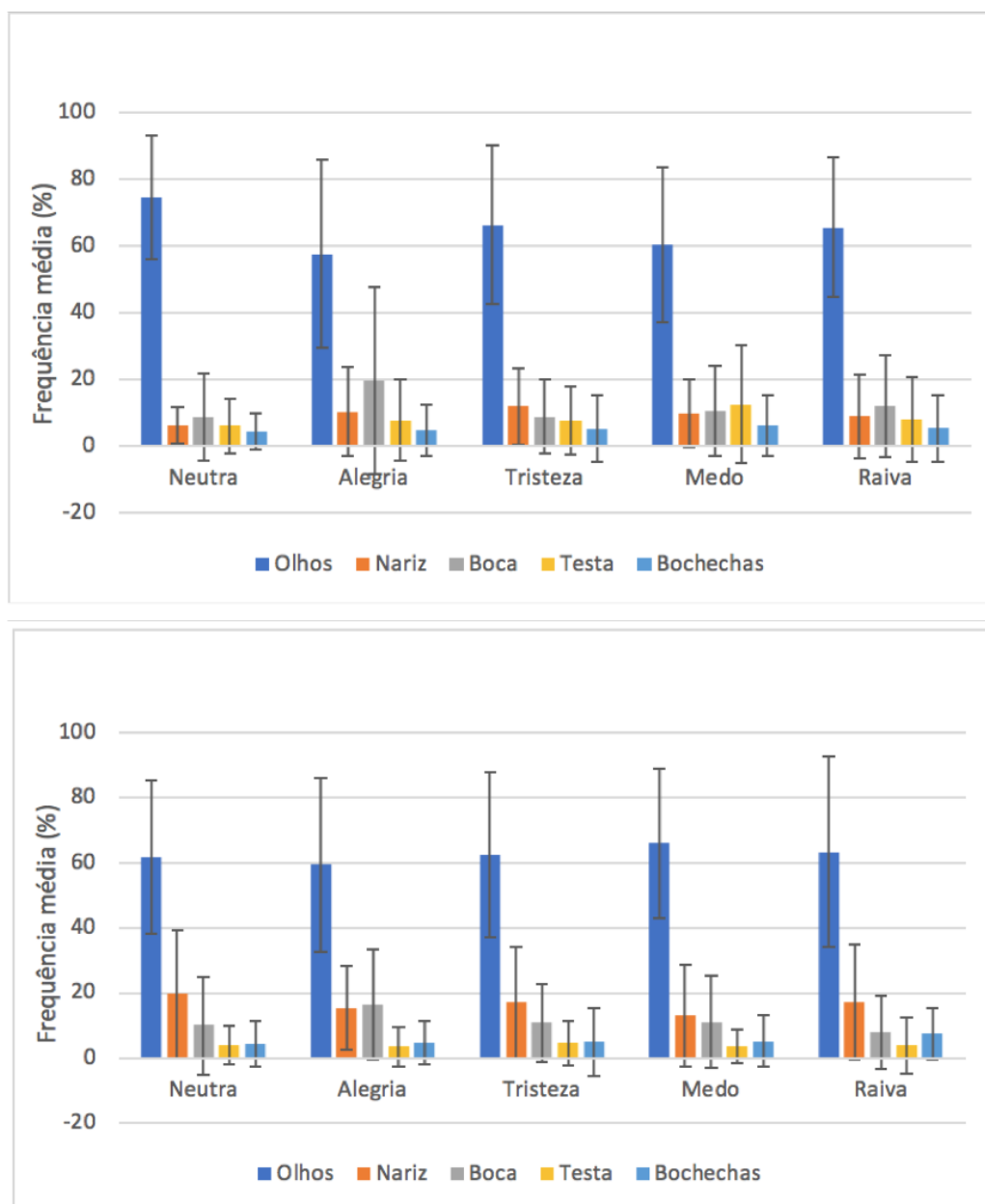


Figura 5.2: Valores médios da frequência da primeira fixação para cada área de interesse relativos às diferentes emoções. Em cima, participantes com formação musical e em baixo, participantes sem formação musical.

Tabela 5.1: Análise de variância - Localização da primeira fixação.

	Com formação musical	Sem formação musical
Fator emoção	não teve um efeito significativo	não teve um efeito significativo
Fator ADI	$F(4, 56) = 48.08$	$F(4, 56) = 36.84$
Interação Emoção*ADI	$F(16, 224) = 2.20$	não teve um efeito significativo

Tabela 5.2: Teste-t - Localização da primeira fixação.

	Com formação musical	Sem formação musical
olhos-testa	$T(14) = 8.45$	$T(14) = 9.92$
olhos-nariz	$T(14) = 8.24$	$T(14) = 4.87$
olhos-bochecha	$T(14) = 9.35$	$T(14) = 8.25$
olhos-boca	$T(14) = 6.74$	$T(14) = 5.86$
boca-testa	não teve um efeito significativo	não teve um efeito significativo
boca-bochechas	não teve um efeito significativo	não teve um efeito significativo

Os indivíduos com formação musical apresentaram uma maior frequência da primeira fixação nos olhos em relação aos participantes sem este tipo de formação, exceto no caso da emoção Alegria. Para além disto, no grupo com formação musical, os olhos são significativamente mais vezes o alvo da primeira fixação para todas as emoções exceto na emoção Alegria. O mesmo não acontece no grupo sem formação musical, pois apenas nas emoções Medo e Tristeza os olhos têm uma frequência significativamente maior em relação às restantes áreas de interesse.

Apesar de, tanto o grupo de indivíduos com formação musical como o grupo sem formação musical ter apenas 15 indivíduos cada, e o tamanho desta amostra ser inferior a 30 (tamanho da amostra recomendado para se realizar um estudo estatístico), foi feita uma análise estatística para cada grupo e uma comparação entre os resultados obtidos nestes testes. Na tabela 5.1, é possível observar que o fator área de interesse (ADI), teve um efeito maior no grupo com formação musical em relação ao grupo sem este tipo de formação. A interação entre os fatores emoção e área de interesse apenas foi significativo para o grupo com formação musical. Na tabela 5.2, estão apresentados os testes-t realizados entre as diferentes áreas de interesse, independentemente da emoção apresentada, para estes dois grupos. Quanto maior for o valor de t , maior será a diferença entre as médias respetivas às áreas de interesse.

5.2 Duração média das fixações

A duração média das fixações representa, em percentagem, o tempo de fixação em cada área de interesse em relação ao tempo de fixação total. Esta normalização foi necessária para se poder proceder a uma comparação da importância relativa de cada área de interesse porque, entre indivíduos e até para o mesmo indivíduo, o tempo de fixação total de cada imagem-estímulo não é constante, podendo até apresentar uma grande variabilidade. Deste modo, está-se a avaliar o peso relativo médio que a amostra de participantes dá a cada área de interesse para as diferentes emoções.

Na Figura 5.3, podemos observar que os olhos são a zona da face que é observada durante mais tempo, independentemente do tipo de emoção apresentada. Observou-se também que na emoção Alegria, a região da boca assume uma importância maior quando comparada com a mesma região nas restantes emoções.

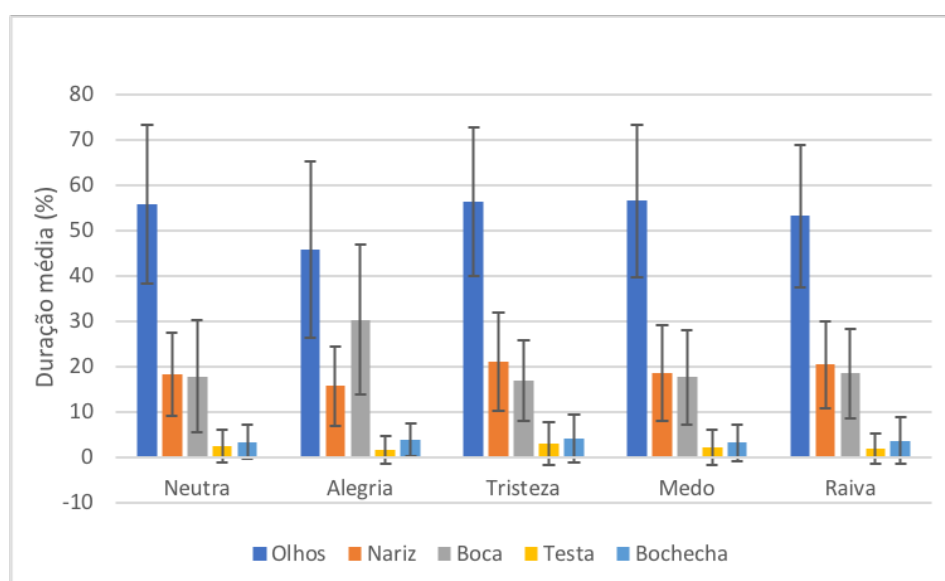


Figura 5.3: Duração média relativa das fixações.

Não se verificou um efeito significativo para o fator emoção. Verificou-se um efeito significativo para o fator área de interesse, $F(4, 116) = 123.48; p < .001; \eta^2_{\text{parcial}} = .81$, e uma interação dos fatores emoção e área de interesse significativo, $F(16, 464) = 10.02; p < .001; \eta^2_{\text{parcial}} = .26$.

Verificou-se que os olhos são alvo de fixações mais longas em relação a todas as restantes áreas da face: testa, $T(29) = 17.90; p < .001$, bochechas $T(29) = 16.31; p < .001$, nariz, $T(29) = 8.62; p < .001$, boca, $T(29) = 7.87; p < .001$. Verificou-se que a boca também é significativamente alvo de atenção em relação à testa, $T(29) = 9.67; p < .001$ e às bochechas, $T(29) = 8.76; p < .001$. Não se verificou um efeito significativo da boca em relação ao nariz.

Tabela 5.3: Análise de variância - Duração média das fixações.

	Com formação musical	Sem formação musical
Fator emoção	não teve um efeito significativo	não teve um efeito significativo
Fator ADI	$F(4, 56) = 80.12$	$F(4, 56) = 37.62$
Interação Emoção*ADI	$F(16, 224) = 6.70$	$F(16, 224) = 3.70$

5.2.1 Validação geral

Estes resultados seguem a mesma tendência que os resultados encontrados em [3], pois foi também encontrado um efeito significativo do fator área de interesse e da interação entre o fator de interesse e a emoção. Também no referido estudo, não se verificou um efeito significativo do fator emoção. Os olhos foram também a região da face onde a atenção visual dos participantes despendeu mais tempo.

5.2.2 Efeito da formação musical

Na Figura 5.4, podemos observar as diferenças entre as durações de fixação médias para cada área de interesse e para cada emoção relativamente aos grupos com e sem formação musical.

O grupo de participantes com formação musical apresentou uma duração de fixação nos olhos superior em todas as emoções quando comparada com os indivíduos sem formação musical. Ambos os grupos apresentam uma duração de fixação nos olhos significativamente superior às restantes zonas da face, exceto no caso da emoção Alegria. O tempo despendido na fixação da região da boca na emoção de Alegria é superior em ambos os grupos quando comparado com as restantes emoções, porém, nos indivíduos com formação musical esta duração é superior em relação aos indivíduos sem formação musical.

Na tabela 5.3, é possível observar que o fator ADI teve um efeito maior no grupo com formação musical em relação ao grupo sem este tipo de formação. A interação entre os fatores emoção e área de interesse apenas foi também maior para o grupo com formação musical. Na tabela 5.4, estão apresentados os testes-t realizados entre as diferentes áreas de interesse, independentemente da emoção apresentada, para estas duas amostras.

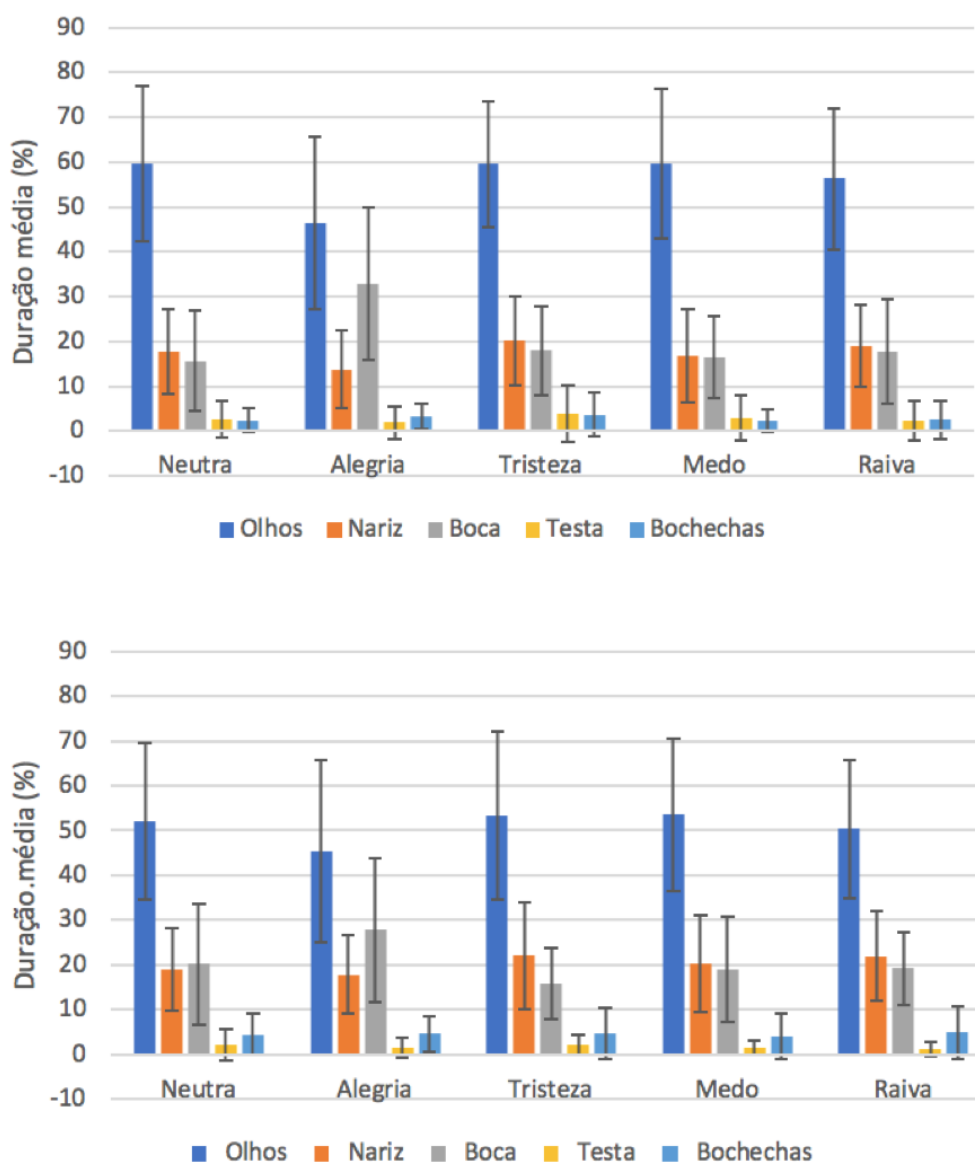


Figura 5.4: Duração média relativa das fixações. Em cima, participantes com formação musical e em baixo, participantes sem formação musical.

Tabela 5.4: Teste-t - Duração média das fixações.

	Com formação musical	Sem formação musical
olhos-testa	T(14)= 13.60	T(14)= 11.61
olhos-nariz	T(14)= 7.41	T(14)= 5.01
olhos-bochecha	T(14)= 14.36	T(14)= 9.64
olhos-boca	T(14)= 6.53	T(14)= 4.69
boca-testa	T(14)= 6.38	T(14)= 7.12
boca-bochechas	T(14)= 7.05	T(14)= 5.36

5.3 Número médio de fixações

O número médio das fixações representa, em percentagem, o número de fixações em cada área de interesse em relação ao número total de fixações. Devido ao facto de o número de fixações não ser igual para todos os participantes nem para todas as imagens-estímulo, pois certos participantes efetuam fixações mais longas, resultando num menor número de fixações, esta métrica necessitou de ser normalizada. Assim, foi possível comparar a frequência média relativa de fixações em cada área de interesse para cada emoção.

No Figura 5.5, é possível observar que os olhos são a zona da face que é fixada mais frequentemente, independentemente do tipo de emoção apresentada. Observou-se também que na emoção Alegria, a região da boca apresenta uma frequência de fixações maior quando comparada com a mesma região nas restantes emoções.

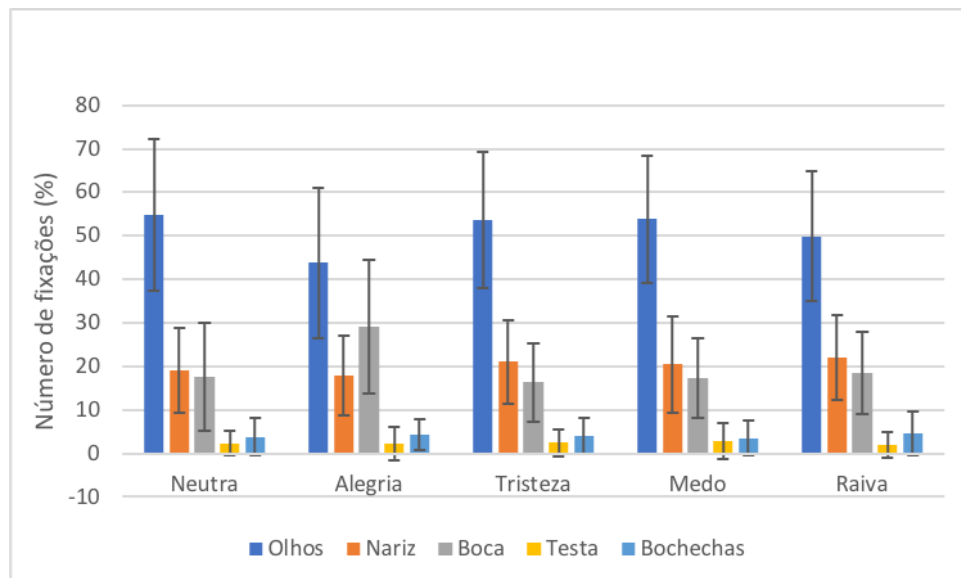


Figura 5.5: Número médio relativo de fixações.

Não se verificou um efeito significativo do fator emoção. O fator área de interesse teve um efeito significativo, $F(4, 166) = 112.01; p < .001; \eta^2_{\text{parcial}} = .81$, e a interação entre o fator emoção e área de interesse teve também um efeito significativo, $F(16, 116) = 9.38; p < .001; \eta^2_{\text{parcial}} = .24$.

Verificou-se um número de fixações significativamente maior nos olhos em relação à testa, $T(29) = 18.21; p < .001$, nos olhos em relação às bochechas, $T(29) = 16.76; p < .001$, nos olhos em relação ao nariz, $T(29) = 8.23; p < .001$, e nos olhos em relação à boca, $T(29) = 8.09; p < .001$. Verificou-se um efeito significativo da boca em relação à testa, $T(29) = 10.07; p < .001$, e da boca em relação às bochechas, $T(29) = 8.62; p < .001$, porém não se verificou um efeito significativo da boca em relação ao nariz.

5.3.1 Validação geral

Estes resultados seguem a mesma tendência que os resultados encontrados em [3], pois foi também encontrado um efeito significativo do fator área de interesse e da interação entre o fator de interesse e a emoção. Também, no referido estudo, não se verificou um efeito significativo do fator emoção. Os olhos foram também fixados mais vezes relativamente a qualquer outra área de interesse.

5.3.2 Efeito da formação musical

Na Figura 5.6, podemos observar as diferenças entre o número de fixações médio para cada área de interesse e para cada emoção relativamente aos grupos com e sem formação musical. A tipologia destes gráficos difere da tipologia dos gráficos anteriores, pois através destes são visualmente mais evidentes as semelhanças e diferenças entre as linhas correspondentes às várias emoções.

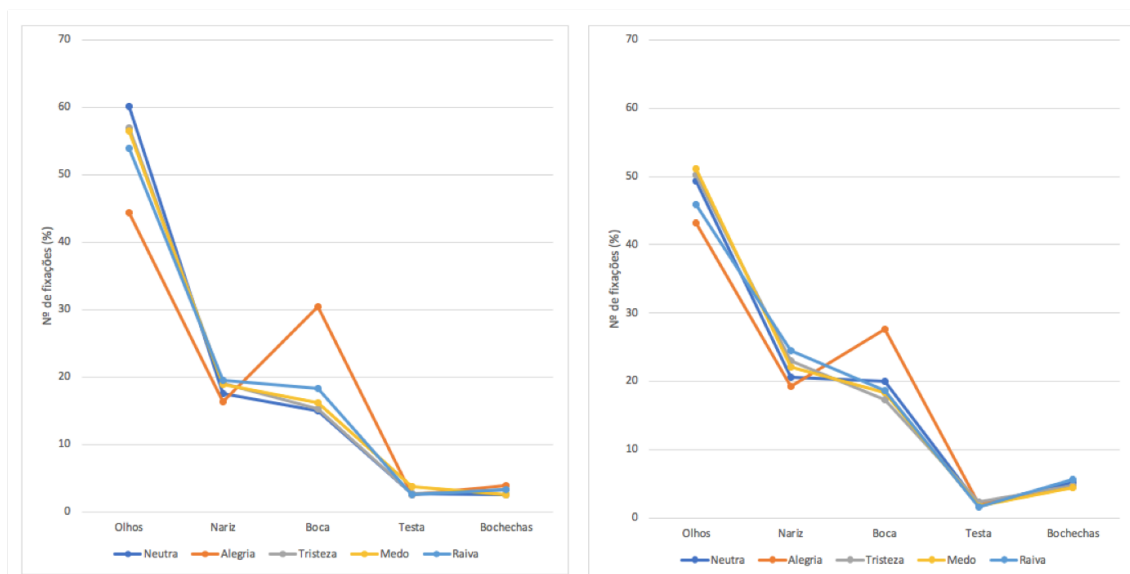


Figura 5.6: Número médio relativo de fixações. À esquerda, participantes com formação musical e à direita, participantes sem formação musical.

O grupo de participantes com formação musical apresentou uma frequência de fixações superior nos olhos relativamente ao grupo de participante sem formação musical. É possível verificar também que, apesar de ambos apresentarem um maior número de fixações na região da boca na emoção Alegria, este valor é mais diferenciado relativamente às restantes emoções nos participantes com formação musical.

Tabela 5.5: Análise de variância - Número médio de fixações.

	Com formação musical	Sem formação musical
Fator emoção	não teve um efeito significativo	não teve um efeito significativo
Fator ADI	$F(4, 56) = 74.82$	$F(4, 56) = 50.65$
Interação Emoção*ADI	$F(16, 224) = 6.71$	$F(16, 224) = 3.36$

Tabela 5.6: Teste-t - Número médio de fixações.

	Com formação musical	Sem formação musical
olhos-testa	$T(14) = 13.62$	$T(14) = 12.16$
olhos-nariz	$T(14) = 7.08$	$T(14) = 4.81$
olhos-bochecha	$T(14) = 14.47$	$T(14) = 10.19$
olhos-boca	$T(14) = 6.91$	$T(14) = 4.71$
boca-testa	$T(14) = 6.38$	$T(14) = 7.79$
boca-bochechas	$T(14) = 6.50$	$T(14) = 5.57$

Na tabela 5.5, é possível observar que o fator ADI teve um efeito maior no grupo com formação musical em relação ao grupo sem este tipo de formação. A interação entre os fatores emoção e área de interesse apenas foi também maior no grupo com formação musical. Na tabela 5.6, estão apresentados os testes-t realizados entre as diferentes áreas de interesse, independentemente da emoção apresentada, para estas duas amostras.

5.4 Rácio olhos-boca

O rácio olhos-boca compara a duração das fixações nos olhos em relação à duração das fixações na boca sendo que, quanto maior for o valor do rácio, maior será a dominância da região dos olhos em relação à região da boca. Deste modo é possível comparar a importância da boca na descodificação da expressão facial emocional.

Na Figura 5.7, observar que a emoção Alegria apresenta um rácio inferior em relação às restantes emoções, demonstrando assim uma dominância dos olhos muito mais baixa.

Verificou-se um efeito significativo para a emoção, $F(4, 145) = 4.77; p = .001; \eta^2_{\text{parcial}} = .26$. Verificou-se um rácio significativamente mais elevado da emoção Neutra relativamente à Alegria, $T(29) = 4.62; p < .001$, da Tristeza em relação à Alegria, $T(29) = 4.76; p < .001$, do Medo em relação à Alegria, $T(29) = 4.16; p < .001$, e da Raiva em relação à Alegria, $T(29) = 4.04; p < .001$.

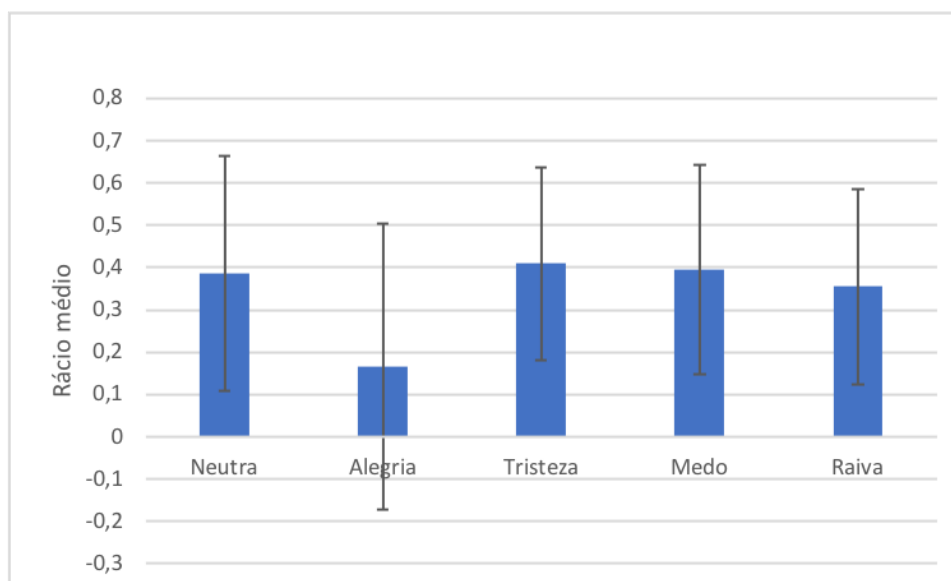


Figura 5.7: Rácio médio olhos-boca ((duração de fixação nos olhos-duração de fixação na boca)/(duração de fixação nos olhos+duração de fixação na boca)).

5.4.1 Validação geral

Os resultados obtidos seguem a mesma tendência dos resultados encontrados em [3] relativamente a todas as emoções, excetuando o caso da emoção Raiva que, no referido estudo, apresenta um valor médio mais baixo do que o valor obtido no presente estudo. Ou seja, a região da boca teve alguma importância na decodificação desta emoção em comparação com a decodificação de outras emoções. Este fenómeno não se verificou de uma forma significativa no presente estudo.

5.4.2 Efeito da formação musical

Na figura 5.8, é possível observar as diferenças entre os rácios correspondentes à população com formação musical e à população sem formação musical.

Os rácios médios dos participantes com formação musical apresentaram valores mais elevados relativamente aos rácios dos participantes sem este tipo de formação para as emoções: Neutra, Tristeza, Medo e Raiva. Na emoção Alegria os participantes com formação musical apresentaram um valor mais baixo para este rácio.

Na tabela 5.7, é possível observar que o fator emoção teve um efeito significativo para o grupo com formação musical, porém este fator não teve um efeito significativo para os indivíduos sem este tipo de formação. Na tabela 5.8, estão apresentados os testes-t realizados entre a emoção Alegria e as restantes emoções, para estes dois grupos.

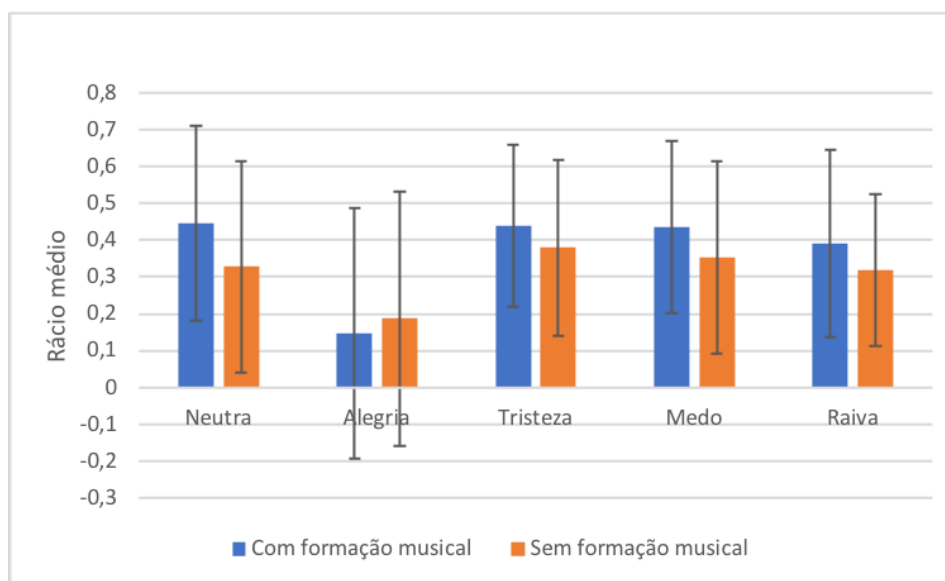


Figura 5.8: Rácio médio olhos-boca ((duração de fixação nos olhos-duração de fixação na boca)/(duração de fixação nos olhos+duração de fixação na boca), relativo a participantes com e sem formação musical

Tabela 5.7: Análise de variância - Rácio olhos-boca.

	Com formação musical	Sem formação musical
Fator emoção	$F(4,70) = 4.50$	não teve um efeito significativo

Tabela 5.8: Teste-t - Rácio olhos - boca.

	Com formação musical	Sem formação musical
Alegria-Neutra	$T(14) = 4.39$	$T(14) = 3.69$
Alegria-Medo	$T(14) = 3.54$	$T(14) = 2.96$
Alegria-Tristeza	$T(14) = 3.46$	$T(14) = 3.06$
Alegria-Raiva	$T(14) = 3.75$	não teve um efeito significativo

5.5 Distância total

A distância total representa a distância na imagem percorrida pelos olhos na imagem-estímulo, ou seja, é a soma das distâncias entre os vários pontos de fixação. Esta métrica é utilizada para se poder estudar se certas emoções provocam padrões de olhar mais dispersos ou mais concentrados consoante a emoção apresentada.

Na Figura 5.9, podemos observar de que forma a distância percorrida varia em função da emoção.

Apesar de existirem diferenças entre as distâncias nos diferentes tipos de emoção, esta variação não se verificou como significativa relativamente ao fator emoção, podendo-se

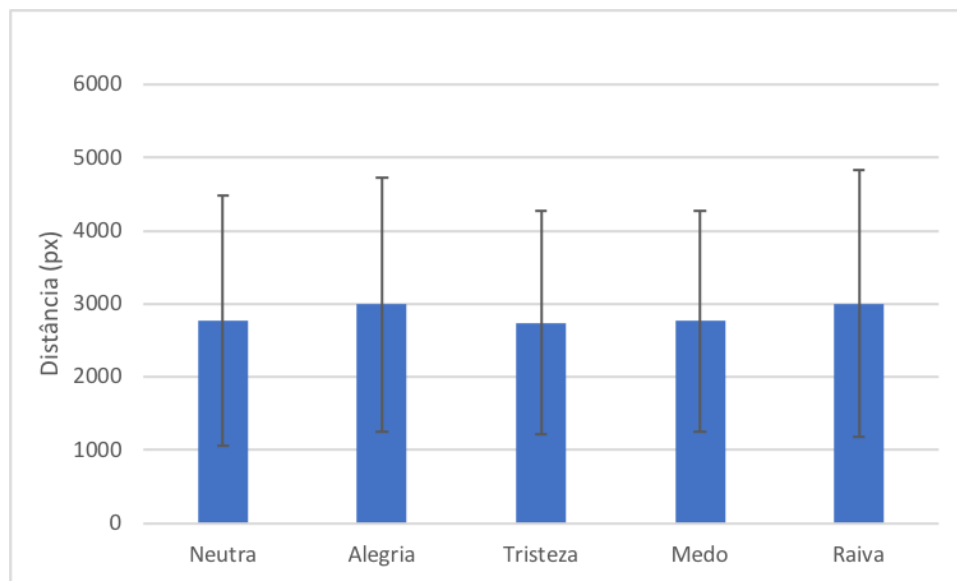


Figura 5.9: Distância média percorrida pelo olhar.

concluir que o tipo de emoção apresentada não influencia significativamente a distância percorrida pelos olhos na imagem estímulo.

5.5.1 Validação geral

Esta métrica não foi aplicada em [3], portanto não existem dados comparativos para validar este resultado.

5.5.2 Efeito da formação musical

Na Figura 5.10, é possível observar as diferenças entre as distâncias correspondentes à população com formação musical e sem formação musical.

As distâncias médias do grupo de participantes com formação musical apresenta um valor mais baixo para todas as emoções em relação ao grupo de participantes sem formação musical, apesar desta diferença não ter sido significativa.

5.6 Discussão

As métricas que foram aplicadas aos dados obtidos através do dispositivo de *eye tracking* permitiram estudar a atenção visual dos indivíduos durante o processo de decodificação de expressões faciais emocionais.

De um modo geral, os olhos tiveram uma importância significativa neste processo. Para além de terem sido o alvo mais frequente da primeira fixação, os olhos foram também fixados durante mais tempo e mais frequentemente, independentemente da emoção apresentada. As regiões do nariz e da boca também tiveram uma importância significativa

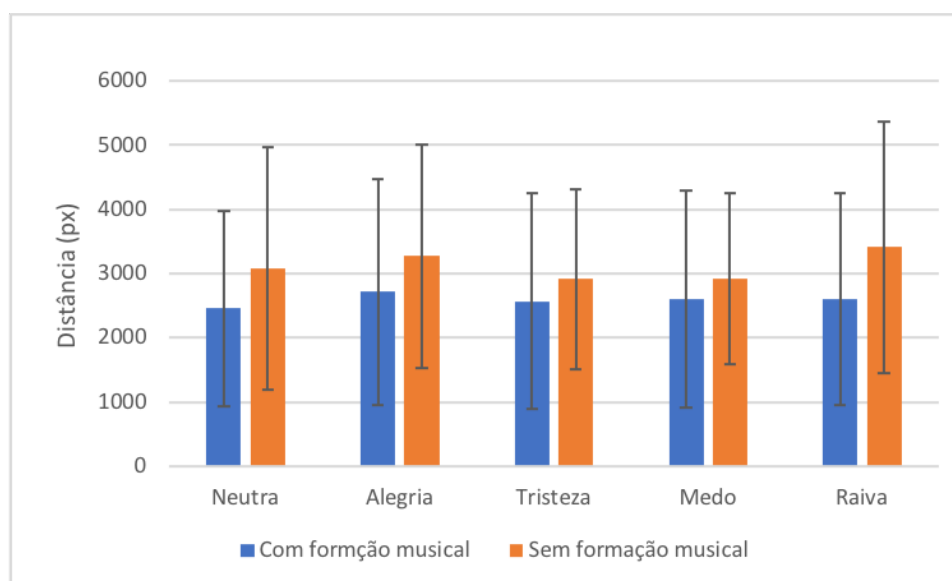


Figura 5.10: Distância média percorrida pelo olhar, relativa a participantes com e sem formação musical.

relativamente às regiões da testa e das bochechas. A atenção visual dada ao nariz pode-se dever ao facto de este estar situado na zona central da face e assim, através da visão periférica, fornecer ao observador informação sobre a expressão facial como um todo. A boca também foi importante neste processo de decodificação porque contém informação relevante, principalmente na emoção Alegria. As regiões das bochechas e da testa foram significativamente fixadas durante menos tempo e com menor frequência, podendo isto significar que contém uma menor quantidade de informação relativamente ao tipo de emoção apresentada. Observou-se que, apesar de a região dos olhos ser dominante em relação à região da boca, independentemente do tipo de emoção apresentada, na emoção Alegria, a boca assume uma importância maior fazendo com que o rácio olhos-boca seja relativamente menor. A distância percorrida pelo olhar durante este processo não foi afetada significativamente pela emoção apresentada apesar de, na emoção Alegria e Raiva, esta distância ter sido ligeiramente maior em relação às restantes emoções. Estas duas emoções apresentaram um rácio olhos-boca menor, portanto o facto da distância percorrida ter apresentado um valor maior pode ser resultado de a atenção visual não permanecer tanto na região dos olhos tendo os indivíduos explorado outras zonas da face, nomeadamente as regiões da boca e do nariz.

Apesar de o grupo com formação musical e o grupo sem formação musical terem apresentado uma tendência semelhante relativamente à dinâmica visual durante a decodificação das expressões faciais, existiram algumas diferenças entre estes dois grupos. Considerando que as regiões dos olhos e da boca contém mais informação relativamente à emoção representada na expressão facial, o facto de haver tanto uma primeira fixação mais frequente nos olhos bem como uma maior duração e frequência de fixações nesta

mesma região por parte dos indivíduos com formação musical, pode significar que estes indivíduos apresentam uma maior eficiência da descodificação da emoção. De igual modo, os indivíduos com formação musical apresentaram uma maior duração e frequência de fixações na região da boca, comparativamente aos indivíduos sem este tipo de formação, tendo sido este fenómeno mais evidente na emoção Alegria. Foi possível observar que a região das bochechas é fixada com maior frequência no grupo de participantes sem formação musical. Considerando que esta zona da face contém pouca informação acerca do tipo de emoção expressa, este comportamento pode revelar uma descodificação da face menos eficaz. A distância percorrida pelo olhar durante a descodificação das expressões faciais foi menor para o grupo com formação musical. Apesar desta diferença não ter sido significativa, poderá ser indicativa de que indivíduos com formação musical têm uma dinâmica visual menos dispersa, enquanto que os indivíduos sem este tipo de formação podem demonstrar uma capacidade de foco menor e uma exploração da face mais errática.

CONCLUSÃO

Este trabalho teve como objetivo estudar o comportamento diferenciado da atenção visual de indivíduos com e sem formação musical aquando da descodificação de expressões faciais com diferentes cargas emocionais. Para a realização do estudo da atenção visual recorreu-se a um *eye tracker* de baixo custo desenvolvido e otimizado em teses anteriores [1, 2] para registar os movimentos oculares dos participantes e posteriormente se proceder à análise da sua dinâmica visual.

Uma vez que este dispositivo nunca foi utilizado em contexto real, foi necessário proceder à sua validação. Para tal, foi seguida a metodologia de um estudo realizado em 2011, com o objetivo de avaliar a atenção visual de sujeitos, em relação a imagens de expressões faciais com diferentes cargas emocionais [3]. Os resultados obtidos nesta dissertação seguiram uma tendência semelhante aos do referido estudo, tendo-se obtido resultados suficientemente próximos para se assumir a validade do sistema de *eye tracking*. Tal como no referido estudo, foi verificada uma preponderância na atenção dada aos olhos da expressão facial no processo de descodificação da mesma, sendo que estes foram o alvo mais frequente da primeira fixação. Foram também os olhos a área da face com um maior número de fixações e maior duração das mesmas, independentemente da emoção apresentada. A boca e o nariz foram zonas que também apresentaram uma duração e frequência de fixações elevada relativamente às bochechas e à testa. Uma possível explicação para as fixações localizadas no nariz pode ser o facto de este estar localizado no centro da cara e permitir ao observador, através da visão periférica ter uma visão global da expressão facial como um todo. A região da boca foi a terceira região mais fixada, pois contém informação importante da descodificação da emoção e demonstrou ser um alvo mais frequente da atenção visual que o nariz para as expressões faciais de alegria.

Após a validação do dispositivo, foi possível estudar as diferenças do comportamento visual entre os participantes com e sem formação musical. Os participantes com formação musical apresentaram uma primeira localização mais frequente na região dos olhos, bem como uma duração de fixações e frequência de fixações mais elevada nesta região comparativamente aos participantes sem formação musical. Na expressão facial de Alegria verificou-se também uma duração e frequência de fixações na boca superior no grupo com formação musical relativamente ao grupo sem este tipo de formação. A dominância da região dos olhos relativamente à região da boca foi maior nos indivíduos com formação musical e, esta dominância verificou-se para todas as emoções exceto para a emoção alegria. A distância percorrida pelo olhar foi inferior no grupo de indivíduos com formação musical.

As diferenças apresentadas entre os indivíduos com e sem formação musical devem-se possivelmente a uma sensibilidade emocional maior por parte dos indivíduos com formação musical, demonstrando um foco da atenção visual nas regiões da face que fornecem mais informação sobre a emoção expressa. Este comportamento visual poderá resultar numa maior eficiência na decodificação das expressões faciais emocionais apresentadas.

No decorrer deste estudo surgiram algumas dificuldades que tornaram o processo de análise dos resultados mais moroso e complexo do que era inicialmente esperado. Os testes anteriores realizados com o dispositivo de *eye tracking* de baixo tiveram uma duração de cerca de 5 vezes menor do que a duração dos testes do presente estudo. Consequentemente, a posição do dispositivo em relação aos olhos nem sempre se manteve constante o que resultou, em alguns casos, num desvio da estimativa da localização das coordenadas referentes à atenção visual do participante.

6.1 Trabalho futuro

Como trabalho futuro sugiro que se altere o suporte de cartão para que este se torne mais confortável para o participante durante estudos de duração mais longa, como é o caso do presente estudo. A forma como o dispositivo assenta na face terá que ser repensada para evitar uma variação da posição do dispositivo e um consequente desvio da localização coordenadas. Verificou-se que, nos participantes que possuem uma cana do nariz mais fina, o dispositivo assenta nas maçãs do rosto ficando assim mais suscetível a variações graduais e acidentais resultantes de movimentos faciais e da pressão exercida pelo dispositivo na face. Para colmatar esta situação, seria importante alterar o suporte para o nariz de modo a que este seja adaptável às diferentes anatomias faciais.

BIBLIOGRAFIA

- [1] F. Gamas. “Desenvolvimento de um Eye Tracker de baixo custo”. Tese de mestrado. FCT-UNL, 2017.
- [2] R. Vences. “Validação e calibração de um Eye Tracker de baixo custo”. Tese de mestrado. FCT-UNL, 2018.
- [3] H. Eisenbarth e G. W. Alpers. “Happy mouth and sad eyes: scanning emotional facial expressions.” Em: *Emotion* 11.4 (2011), pp. 860–865.
- [4] L. Pessoa, M. McKenna, E. Gutierrez e L. G. Ungerleider. “Neural processing of emotional faces requires attention”. Em: *Proceedings of the National Academy of Sciences* 99.17 (2002), pp. 11458–11463.
- [5] B. Duchaine e G. Yovel. “A revised neural framework for face processing”. Em: *Annual Review of Vision Science* 1 (2015), pp. 393–416.
- [6] K. E. Stephan, J. C. Marshall, K. J. Friston, J. B. Rowe, A. Ritzl, K. Zilles e G. R. Fink. “Lateralized cognitive processes and lateralized task control in the human brain”. Em: *Science* 301.5631 (2003), pp. 384–386.
- [7] K. Ono, A. Nakamura, K. Yoshiyama, T. Kinkori, M. Bundo, T. Kato e K. Ito. “The effect of musical experience on hemispheric lateralization in musical feature processing”. Em: *Neuroscience letters* 496.2 (2011), pp. 141–145.
- [8] E. Smith e S. Kosslyn. *Cognitive Psychology: Mind and Brain (International Edition)*. 2007.
- [9] N. Gosselin, I. Peretz, D. Hasboun, M. Baulac e S. Samson. “Impaired recognition of musical emotions and facial expressions following anteromedial temporal lobe excision”. Em: *Cortex* 47.9 (2011), pp. 1116–1125.
- [10] A. D. V. Sabino, C. M. Camargo, M. H. N. Chagas e F. L. Osório. “Facial recognition of happiness is impaired in musicians with high music performance anxiety”. Em: *Frontiers in psychiatry* 9 (2018), pp. 5–9.
- [11] P. T. R. Seeley T. Stephens. “Anatomy and Fisiology”. Em: McGraw-Hill, 2008. Cap. 15, pp. 536–557.
- [12] O. G. Sebastian Ochlenburg. “The lateralized brain-The neuroscience and evolution of hemispheric asymmetries”. Em: Academic Press, 2018. Cap. 1, pp. 1–25.

- [13] R. S. Aisami. "Learning styles and visual literacy for learning and performance". Em: *Procedia-Social and Behavioral Sciences* 176 (2015), pp. 538–545.
- [14] O. G. Sebastian Ochlenbrugg. "The lateralized brain-The neuroscience and evolution of hemispheric asymmetries". Em: Academic Press, 2018. Cap. 4, pp. 87–121.
- [15] O. G. Sebastian Ochlenbrugg. "The lateralized brain-The neuroscience and evolution of hemispheric asymmetries". Em: Academic Press, 2018. Cap. 7, pp. 185–238.
- [16] J. Pickles. *An introduction to the physiology of hearing*. Brill, 2013.
- [17] N. Kanwisher, J. McDermott e M. M. Chun. "The fusiform face area: a module in human extrastriate cortex specialized for face perception". Em: *Journal of neuroscience* 17.11 (1997), pp. 4302–4311.
- [18] D. V. Buonomano e M. M. Merzenich. "Cortical plasticity: from synapses to maps". Em: *Annual review of neuroscience* 21.1 (1998), pp. 149–186.
- [19] C. Gaser e G. Schlaug. "Brain structures differ between musicians and non-musicians". Em: *Journal of Neuroscience* 23.27 (2003), pp. 9240–9245.
- [20] G. Schlaug, L. Jäncke, Y. Huang, J. F. Staiger e H. Steinmetz. "Increased corpus callosum size in musicians". Em: *Neuropsychologia* 33.8 (1995), pp. 1047–1055.
- [21] J. Sergent. "Mapping the musician brain". Em: *Human Brain Mapping* 1.1 (1993), pp. 20–38.
- [22] D. Hebb. "Brain mechanisms and learning". Em: *Distinctive features of learning in the higher animal* (1961), pp. 37–46.
- [23] C. Hou e Z. Liu. "The Survival Processing Advantage of Face: The Memorization of the (Un) Trustworthy Face Contributes More to Survival Adaptation." Em: *Evolutionary Psychology* 17.2 (2019), pp. 49–53.
- [24] P. Vuilleumier, J. L. Armony, J. Driver e R. J. Dolan. "Effects of attention and emotion on face processing in the human brain: an event-related fMRI study". Em: *Neuron* 30.3 (2001), pp. 829–841.
- [25] M. Catani, F. Dell'Acqua e M. T. De Schotten. "A revised limbic system model for memory, emotion and behaviour". Em: *Neuroscience & Biobehavioral Reviews* 37.8 (2013), pp. 1724–1737.
- [26] R. Adolphs, H. Damasio, D. Tranel e A. R. Damasio. "Cortical systems for the recognition of emotion in facial expressions". Em: *Journal of neuroscience* 16.23 (1996), pp. 7678–7687.
- [27] B. Bediou, N. Franck, M. Saoud, J.-Y. Baudouin, G. Tiberghien, J. Daléry e T. d'Amato. "Effects of emotion and identity on facial affect processing in schizophrenia". Em: *Psychiatry research* 133.2-3 (2005), pp. 149–157.

- [28] Z. Boraston, S.-J. Blakemore, R. Chilvers e D. Skuse. "Impaired sadness recognition is linked to social interaction deficit in autism". Em: *Neuropsychologia* 45.7 (2007), pp. 1501–1510.
- [29] H. Eisenbarth, G. W. Alpers, D. Segrè, A. Calogero e A. Angrilli. "Categorization and evaluation of emotional faces in psychopathic women". Em: *Psychiatry research* 159.1-2 (2008), pp. 189–195.
- [30] NYSOA. *Normal Eye Anatomy*. URL: <https://www.nysoa.org/page/normal-eye-anatomy-54.html>. (accessed: 25.01.2020).
- [31] A. Fiorentini. "Differences between fovea and parafovea in visual search processes". Em: *Vision research* 29.9 (1989), pp. 1153–1164.
- [32] A. T. Duchowski. "Eye Tracking Methodology". Em: Springer, 2007. Cap. 2, pp. 51–59.
- [33] J. T. Holladay. "Visual acuity measurements". Em: *Journal of Cataract & Refractive Surgery* 30.2 (2004), pp. 287–290.
- [34] A. Kar e P. Corcoran. "A review and analysis of eye-gaze estimation systems, algorithms and performance evaluation methods in consumer platforms". Em: *IEEE Access* 5 (2017), pp. 16495–16519.
- [35] A. M. Barreto. "Eye tracking como método de investigação aplicado às ciências da comunicação". Em: *Revista Comunicando* 1.1 (2012), pp. 168–186.
- [36] R. Dooley. *Neuromarketing*. URL: <https://www.neurosciencemarketing.com/blog/articles/baby-heat-maps.htm>. (accessed: 24.01.2020).
- [37] C. Tool. *Do marketing that captures attention*. URL: <https://cooltool.com/eyetracking>. (accessed: 24.01.2020).
- [38] T. Technology. *Eye Tracking*. URL: <http://hci.cc.metu.edu.tr/en/eye-tracking>. (accessed: 24.01.2020).
- [39] H Zamani, A Abas e M. Amin. "Eye tracking application on emotion analysis for marketing strategy". Em: *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)* 8.11 (2016), pp. 87–91.
- [40] R. Dodge e T. S. Cline. "The angle velocity of eye movements." Em: *Psychological Review* 8.2 (1901), pp. 145–146.
- [41] A. T. Duchowski. "Eye tracking methodology". Em: *Theory and practice* 328.614 (2007), pp. 2–3.
- [42] P. Green. *Review of eye fixation recording methods and equipment*. Rel. téc. 1992.
- [43] T. Technology. *Dark and bright pupil tracking*. URL: <https://www.tobiipro.com/learn-and-support/learn/eye-tracking-essentials/what-is-dark-and-bright-pupil-tracking/>. (accessed: 24.01.2020).

- [44] J. H. Goldberg e A. M. Wichansky. "Eye tracking in usability evaluation: A practitioner's guide". Em: *the Mind's Eye*. Elsevier, 2003, pp. 493–516.
- [45] N. Wade, B. W. Tatler et al. *The moving tablet of the eye: The origins of modern eye movement research*. Oxford University Press, USA, 2005.
- [46] D. W. Cunningham, M. Kleiner, H. H. Bülthoff e C. Wallraven. "The components of conversational facial expressions". Em: *Proceedings of the 1st Symposium on Applied perception in graphics and visualization*. 2004, pp. 143–150.
- [47] M. Nusseck, D. W. Cunningham, C. Wallraven e H. H. Bülthoff. "The contribution of different facial regions to the recognition of conversational expressions". Em: *Journal of vision* 8.8 (2008), pp. 1–10.
- [48] M. L. Smith, G. W. Cottrell, F. Gosselin e P. G. Schyns. "Transmitting and decoding facial expressions". Em: *Psychological science* 16.3 (2005), pp. 184–189.
- [49] R. Adolphs, F. Gosselin, T. W. Buchanan, D. Tranel, P. Schyns e A. R. Damasio. "A mechanism for impaired fear recognition after amygdala damage". Em: *Nature* 433.7021 (2005), pp. 68–72.
- [50] J. Lee, F. Gosselin, J. K. Wynn e M. F. Green. "How do schizophrenia patients use visual information to decode facial emotion?" Em: *Schizophrenia bulletin* 37.5 (2011), pp. 1001–1008.
- [51] K. Dunlap. "The role of eye-muscles and mouth-muscles in the expression of the emotions." Em: *Genetic Psychology Monographs* (1927).
- [52] C. A. Ruckmick. "A preliminary study of the emotions." Em: *Psychological monographs* 30.3 (1921), pp. 30–32.
- [53] S. Baron-Cohen e Wheelwright. "Is there a "language of the eyes"? Evidence from normal adults, and adults with autism or Asperger syndrome". Em: *Visual cognition* 4.3 (1997), pp. 311–331.
- [54] J. C. Coleman. "Facial expressions of emotion." Em: *Psychological Monographs: General and Applied* 63.1 (1949), pp. 33–36.
- [55] R. E. Jack, C. Blais, C. Scheepers, P. G. Schyns e R. Caldara. "Cultural confusions show that facial expressions are not universal". Em: *Current biology* 19.18 (2009), pp. 1543–1548.
- [56] J. N. Bassili. "Emotion recognition: The role of facial movement and the relative importance of upper and lower areas of the face." Em: *Journal of personality and social psychology* 37.11 (1979), pp. 2049–2058.
- [57] R. G. Ley e M. P. Bryden. "Hemispheric differences in processing emotions and faces". Em: *Brain and language* 7.1 (1979), pp. 127–138.
- [58] J. C. Borod, F. Andelman, L. K. Obler, J. R. Tweedy e J. Wilkowitz. "Right hemisphere specialization for the identification of emotional words and sentences: Evidence from stroke patients". Em: *Neuropsychologia* 30.9 (1992), pp. 827–844.

- [59] J. W. Van Strien e M. Morpurgo. "Opposite hemispheric activations as a result of emotionally threatening and non-threatening words". Em: *Neuropsychologia* 30.9 (1992), pp. 845–848.
- [60] S. T. Dekosky, K. M. Heilman, D. Bowers e E. Valenstein. "Recognition and discrimination of emotional faces and pictures". Em: *Brain and language* 9.2 (1980), pp. 206–214.
- [61] I. Fried, C. MATEER, G. Ojemann, R. Wohms e P. Fedio. "Organization of visuospatial functions in human cortex: evidence from electrical stimulation". Em: *Brain* 105.2 (1982), pp. 349–371.
- [62] D. Lundqvist. *KDEF and AKDEF*. URL: <http://kdef.se/home/aboutKDEF.html>. (accessed: 28.01.2020).
- [63] M. M. Bradley e P. J. Lang. "Measuring emotion: the self-assessment manikin and the semantic differential". Em: *Journal of behavior therapy and experimental psychiatry* 25.1 (1994), pp. 49–59.

